



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 171833

**TO: Ginny Portner
Location: 3b02 / 3c18
Tuesday, November 29, 2005
Art Unit: 1645
Phone: 571-272-0862
Serial Number: 09 / 904994**

**From: Jan Delaval
Location: Biotech-Chem Library
Remsen 1a51
Phone: 571-272-2504

jan.delaval@uspto.gov**

Search Notes

STIC-Biotech/ChemLib

171833

From: Portner, Ginny
Sent: Wednesday, November 16, 2005 6:32 PM
To: STIC-Biotech/ChemLib
Subject: 09/904,994

please search SEQ ID NO 1, 2 and 3. thanks

Ginny Portner
Remsen Building
Art Unit 1645
Room E03, B02
(571) 272-0862

3C18

Scan into EDAN
RECEIVED
NOV 17 2005
STIC/CHEN LIBRARY

Jan
22604

u/ralok



STIC SEARCH RESULTS FEEDBACK FORM

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact:*

Mary Hale, Information Branch Supervisor
22507, Remsen 1d86

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC/Biotech-Chem Library CM1 - Circ. Desk



November 2005

Published_Applications Nucleic Acid and Published_Applications Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New).

Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 23:29:19 ; Search time 14055 Seconds
(without alignments)
11659.884 Million cell updates/sec

Title: US-09-904-994B-1
Perfect score: 2883
Sequence: 1 rrgagattttccarctt.....aaaaagtagaacacagg 2883

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_in.*
- 3: gb_env.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pr.*
- 9: gb_ro.*
- 10: gb_stc.*
- 11: gb_sy.*
- 12: gb_un.*
- 13: gb_vi.*
- 14: gb_htg.*
- 15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB | ID | Description |
|------------|--------|---------------|--------|----|------------|---------------------|
| 1 | 2880.6 | 99.9 | 2883 | 6 | BD185302 | BD185302 Helicobac |
| 2 | 2880.6 | 99.9 | 2883 | 6 | AX356683 | AX356683 Sequence |
| 3 | 2228 | 77.3 | 2452 | 6 | BD185306 | BD185306 Helicobac |
| 4 | 2228 | 77.3 | 2452 | 6 | AX356695 | AX356695 Sequence |
| 5 | 2158.6 | 74.9 | 2405 | 6 | BD185303 | BD185303 Helicobac |
| 6 | 2158.6 | 74.9 | 2405 | 6 | AX356686 | AX356686 Sequence |
| 7 | 2158.6 | 74.9 | 2407 | 6 | BD185305 | BD185305 Helicobac |
| 8 | 2158.6 | 74.9 | 2407 | 6 | AX356692 | AX356692 Sequence |
| 9 | 1948 | 67.6 | 2183 | 6 | BD185304 | BD185304 Helicobac |
| 10 | 1948 | 67.6 | 2183 | 6 | AX356689 | AX356689 Sequence |
| 11 | 1185.6 | 41.1 | 8406 | 1 | AF330621 | AF330621 Helicobac |
| 12 | 1138 | 39.5 | 2664 | 1 | HECUREASE | L25079 Helicobacter |
| c 13 | 1090.8 | 37.8 | 12037 | 1 | AE001446 | AE001446 Helicobac |
| c 14 | 1081.6 | 37.5 | 10903 | 1 | AE000529 | AE000529 Helicobac |
| 15 | 1075.2 | 37.3 | 2735 | 6 | AR054309 | AR054309 Sequence |
| 16 | 1075.2 | 37.3 | 5100 | 1 | HECUREABCD | M60398 H.pylori ur |
| 17 | 1072.4 | 37.2 | 7088 | 1 | AB032429 | AB032429 Helicobac |
| 18 | 1062.6 | 36.9 | 2619 | 6 | A41006 | A41006 Sequence 1 |

| | | | | | | |
|------|--------|------|--------|---|-------------|--------------------|
| 19 | 1062.6 | 36.9 | 2619 | 6 | A72509 | A72509 Sequence 1 |
| 20 | 1062.6 | 36.9 | 2619 | 6 | A73285 | A73285 Sequence 1 |
| 21 | 1062.6 | 36.9 | 2619 | 6 | AR061045 | AR061045 Sequence |
| 22 | 1062.6 | 36.9 | 2619 | 6 | AR161678 | AR161678 Sequence |
| 23 | 1060.4 | 36.8 | 2430 | 1 | AY227442 | AY227442 Helicobac |
| 24 | 1058.6 | 36.7 | 4824 | 6 | AR349305 | AR349305 Sequence |
| 25 | 1058.6 | 36.7 | 4824 | 6 | AR349306 | AR349306 Sequence |
| 26 | 1057.8 | 36.7 | 2619 | 1 | HFURE | X69080 H.felis ure |
| 27 | 1056.8 | 36.7 | 2746 | 1 | HPHPUABG | X11709 Helicobacte |
| 28 | 1056.8 | 36.7 | 2767 | 6 | A31515 | A31515 H.pylori ur |
| 29 | 985.2 | 34.2 | 6126 | 1 | AF332656 | AF332656 Helicobac |
| 30 | 985.2 | 34.2 | 30929 | 1 | AE017145 | AE017145 Helicobac |
| 31 | 894.6 | 31.0 | 1704 | 1 | AY714224 | AY714224 Helicobac |
| 32 | 885.8 | 30.7 | 1815 | 6 | BD092530 | BD092530 Identific |
| 33 | 884.8 | 30.7 | 1710 | 1 | AY295085 | AY295085 Helicobac |
| 34 | 881.8 | 30.6 | 1707 | 6 | CQ903244 | CQ903244 Sequence |
| 35 | 881.8 | 30.6 | 1707 | 6 | CQ903245 | CQ903245 Sequence |
| 36 | 880 | 30.5 | 1710 | 1 | AF352376 | AF352376 Helicobac |
| 37 | 878.4 | 30.5 | 1710 | 6 | A07398 | A07398 C.jejuni ge |
| 38 | 878.4 | 30.5 | 1710 | 6 | A08818 | A08818 C.jejuni DN |
| 39 | 876 | 30.4 | 1680 | 1 | AF479026 | AF479026 Helicobac |
| 40 | 772.4 | 26.8 | 110000 | 1 | BA000004_02 | Continuation (3 of |
| 41 | 766 | 26.6 | 10118 | 1 | BSUEROP | Y08559 Bacillus su |
| c 42 | 766 | 26.6 | 200690 | 1 | BSUB0019 | Z99122 Bacillus su |
| 43 | 739.4 | 25.6 | 1665 | 1 | AF508003 | AF508003 Helicobac |
| 44 | 737.6 | 25.6 | 6537 | 1 | BACUREA | D14439 Thermophili |
| 45 | 733.8 | 25.5 | 4723 | 1 | AB201709 | AB201709 Campyloba |

ALIGNMENTS

| | | | | | | |
|------------|---|-----------------------------|---------|-----|--------|-----------------|
| RESULT 1 | BD185302 | BD185302 | 2883 bp | DNA | linear | PAT 17-JUN-2003 |
| LOCUS | BD185302 | Helicobacter felis vaccine. | | | | |
| DEFINITION | BD185302 | Helicobacter felis vaccine. | | | | |
| ACCESSION | BD185302 | | | | | |
| VERSION | BD185302.1 | GI:31877502 | | | | |
| KEYWORDS | JP 2002355054-A/1. | | | | | |
| SOURCE | Helicobacter felis | | | | | |
| ORGANISM | Helicobacter felis | | | | | |
| REFERENCE | Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Helicobacter. | | | | | |
| AUTHORS | 1 (bases 1 to 2883) | | | | | |
| TITLE | Kusters, J.G. and Cattoli, G. | | | | | |
| JOURNAL | Helicobacter felis vaccine | | | | | |
| COMMENT | Patent: JP 2002355054-A 1 10-DEC-2002; | | | | | |
| | AKZO NOBEL NV | | | | | |
| | OS Helicobacter felis | | | | | |
| | PN JP 2002355054-A/1 | | | | | |
| | PD 10-DEC-2002 | | | | | |
| | PF 16-JUL-2001 JP 2001214711 | | | | | |
| | PI 17-JUL-2000 EP 00202565.8 | | | | | |
| | PI JOHANNES GERARDUS KUSTERS, GIOVANNI CATTOLI | | | | | |
| | PC | | | | | |
| | C12N15/09, A61K38/00, A61K39/106, A61K39/118, A61K39/12, A61K39/175, PC | | | | | |
| | A61K39/23, | | | | | |
| | PC A61K39/235, A61K39/39, A61K39/395, A61P1/04, A61P31/04, C12N1/15, | | | | | |
| | PC C12N1/19, | | | | | |
| | PC | | | | | |
| | C12N1/21, C12N5/10, C12N9/80, C12Q1/68, G01N33/15, G01N33/50, G01N33/ PC | | | | | |
| | 53, | | | | | |
| | PC | | | | | |
| | G01N33/53, G01N33/566, G01N33/569// (C12N9/80, C12R1:01), (C12Q1/68, PC | | | | | |
| | C12R1:01), | | | | | |
| | PC C12N15/00, C12N5/00, A61K37/02 | | | | | |
| | CC Helicobacter felis vaccine | | | | | |
| | FT Key Location/Qualifiers | | | | | |
| | FT CDS (206) . . (886) | | | | | |
| | FT CDS (897) . . (2603). | | | | | |
| FEATURES | Location/Qualifiers | | | | | |
| source | 1..2883 | | | | | |
| | /organism="Helicobacter felis" | | | | | |

| | | | | | | | | | |
|--|-----|---|------|--|--|--|--|--|--|
| /mol_type="genomic DNA" /db_xref="taxon:214" | | | | | | | | | |
| ORIGIN | | | | | | | | | |
| Query Match 99.9%; Score 2880.6; DB 6; Length 2883; Best Local Similarity 100.0%; Pred. No. 0; Matches 2883; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | | | | | | | |
| QY | 1 | RGRAGATTTTCCARCACTTCAAGCACATATTGATCCTGTGTTGGTGGGTAAATTCR | 60 | | | | | | |
| DB | 1 | RGRAGATTTTCCARCACTTCAAGCACATATTGATCCTGTGTTGGTGGTAAATTCR | 60 | | | | | | |
| QY | 61 | ACTGTGTAATCTATTAATTAATTTTAAATAATTAATTAATTAATTAATTAATTA | 120 | | | | | | |
| DB | 61 | ACTGTGTAATCTATTAATTAATTTTAAATAATTAATTAATTAATTAATTAATTA | 120 | | | | | | |
| QY | 121 | TTACTTATTAATAAAGTTAATAAAGTAACGAATTAAGGACTATAATCCCATTCGCTT | 180 | | | | | | |
| DB | 121 | TTACTTATTAATAAAGTTAATAAAGTAACGAATTAAGGACTATAATCCCATTCGCTT | 180 | | | | | | |
| QY | 181 | TAAAAATTTAACAAAGAGTAATAGGTGAAACTCACACCAAAAGACGAAGAAAGTTCTT | 240 | | | | | | |
| DB | 181 | TAAAAATTTAACAAAGAGTAATAGGTGAAACTCACACCAAAAGACGAAGAAAGTTCTT | 240 | | | | | | |
| QY | 241 | GTTATATTATGCGGGCGAAGTGCTAGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCA | 300 | | | | | | |
| DB | 241 | GTTATATTATGCGGGCGAAGTGCTAGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCA | 300 | | | | | | |
| QY | 301 | ACCGNAGCATGCTTACATTAGTCCCATATTATGGACGAAGCGCGCGTGGAAAAAA | 360 | | | | | | |
| DB | 301 | ACCGNAGCATGCTTACATTAGTCCCATATTATGGACGAAGCGCGCGTGGAAAAAA | 360 | | | | | | |
| QY | 361 | AACCGTTGCCAGCTTATGGAGAGTGCACTTTTGAAGAAAGATGAAGTAATGCC | 420 | | | | | | |
| DB | 361 | AACCGTTGCCAGCTTATGGAGAGTGCACTTTTGAAGAAAGATGAAGTAATGCC | 420 | | | | | | |
| QY | 421 | CGGGTGGGTAATATGGTTCCCGATCTAGGTGTAAGGCCACTTTCCTGATGGTAGAA | 480 | | | | | | |
| DB | 421 | CGGGTGGGTAATATGGTTCCCGATCTAGGTGTAAGGCCACTTTCCTGATGGTAGAA | 480 | | | | | | |
| QY | 481 | ACTTGTAACTTGAATTTGGCCCATCGAACAGATGAGCACTTCAAAGCGGGCGAAGTGAA | 540 | | | | | | |
| DB | 481 | ACTTGTAACTTGAATTTGGCCCATCGAACAGATGAGCACTTCAAAGCGGGCGAAGTGAA | 540 | | | | | | |
| QY | 541 | ATTTGGTTCGATAAAGACATCGAGCTCAATGCGGCAAGAGTAACCGAATTCAGGT | 600 | | | | | | |
| DB | 541 | ATTTGGTTCGATAAAGACATCGAGCTCAATGCGGCAAGAGTAACCGAATTCAGGT | 600 | | | | | | |
| QY | 601 | TACTAATGAAGGCGCTTAAATCCTTGATGTGGTAGCCATTTCACCTTCTTTGAAAGCTAA | 660 | | | | | | |
| DB | 601 | TACTAATGAAGGCGCTTAAATCCTTGATGTGGTAGCCATTTCACCTTCTTTGAAAGCTAA | 660 | | | | | | |
| QY | 661 | CAAGGCATTAATAATTCGATGTGAAGAAAGCCTATGGCAAAAGCGCTAGATATTCCTCTGG | 720 | | | | | | |
| DB | 661 | CAAGGCATTAATAATTCGATGTGAAGAAAGCCTATGGCAAAAGCGCTAGATATTCCTCTGG | 720 | | | | | | |
| QY | 721 | CAACCGCTACGATTTGGGCGAGCAAAACCCGCAAGTGCAGTTGATTCCTCTTGGTGG | 780 | | | | | | |
| DB | 721 | CAACCGCTACGATTTGGGCGAGCAAAACCCGCAAGTGCAGTTGATTCCTCTTGGTGG | 780 | | | | | | |
| QY | 781 | CAGTAAAAAGTGAATGGGATGAACGGGCTTTGTAATAACATCGCGGATGAACGCCATAA | 840 | | | | | | |
| DB | 781 | CAGTAAAAAGTGAATGGGATGAACGGGCTTTGTAATAACATCGCGGATGAACGCCATAA | 840 | | | | | | |
| QY | 841 | ACATAAGCGCTTGAACAAGCGGAAATCTCAAGGATTTATCAAGTAAGGAGACTCCCATGA | 900 | | | | | | |
| DB | 841 | ACATAAGCGCTTGAACAAGCGGAAATCTCAAGGATTTATCAAGTAAGGAGACTCCCATGA | 900 | | | | | | |
| QY | 901 | AAATGAAAAAACAAGATATGTAAATACCTACCGACCCCAAGGGCGATAAAGTGGCT | 960 | | | | | | |
| DB | 901 | AAATGAAAAAACAAGATATGTAAATACCTACCGACCCCAAGGGCGATAAAGTGGCT | 960 | | | | | | |
| QY | 961 | TAGGAGATACCGATCTTTGGGCGAAGTAGAACATGACTATACCACCTATGGCGAAGAAC | 1020 | | | | | | |

| | | | |
|------------|---|---|----------------------------|
| Db | 2041 | ATATAAAGAAATTTGGTGAAGCTTCCTGAAGATGCCAAGATAACGATAAATTTCCGCATTA | 2100 |
| QY | 2101 | AGCGCTACATCTCCAAATACATATCAACCCCGCTTTGAACCCAGCGGTGAGCGAGTATA | 2160 |
| Db | 2101 | AGCGCTACATCTCCAAATACATATCAACCCCGCTTTGAACCCAGCGGTGAGCGAGTATA | 2160 |
| QY | 2161 | TGCGCTCTGTGGAAGAGGCAAGATCCCGACCTGTGGTGTGGAAATCTCGCCTTTTTTG | 2220 |
| Db | 2161 | TGCGCTCTGTGGAAGAGGCAAGATCCCGACCTGTGGTGTGGAAATCTCGCCTTTTTTG | 2220 |
| QY | 2221 | GGTAAACCCAAATCGTGATCAAGCGCGTATGGTCTTCTCTGAATGGCGCAT | 2280 |
| Db | 2221 | GGTAAACCCAAATCGTGATCAAGCGCGTATGGTCTTCTCTGAATGGCGCAT | 2280 |
| QY | 2281 | CTAACCGCTCTGTGCCACATCCCAACCGGTTTATTACCGGAAATGTTTGGGCATCACG | 2340 |
| Db | 2281 | CTAACCGCTCTGTGCCACATCCCAACCGGTTTATTACCGGAAATGTTTGGGCATCACG | 2340 |
| QY | 2341 | GCAAGCGGAAATTTGACACGAGCATCATTTTGTGTTTCCAAAGTCGCCTATGAATAAGCG | 2400 |
| Db | 2341 | GCAAGCGGAAATTTGACACGAGCATCATTTTGTGTTTCCAAAGTCGCCTATGAATAAGCG | 2400 |
| QY | 2401 | TGAAAGAAAGCTGGGCTTAGAGGCCAAGTTCTACCGGTCAAAAACCTGCGGTAACATCA | 2460 |
| Db | 2401 | TGAAAGAAAGCTGGGCTTAGAGGCCAAGTTCTACCGGTCAAAAACCTGCGGTAACATCA | 2460 |
| QY | 2461 | CCAAGAAAGACTTCAAGTCAACGACAAACCGGCAAAATCACCCTCGATCCGAAAACCT | 2520 |
| Db | 2461 | CCAAGAAAGACTTCAAGTCAACGACAAACCGGCAAAATCACCCTCGATCCGAAAACCT | 2520 |
| QY | 2521 | TCGAGTCTTTGTAGATGGCAAACTCTGCACCTCTAAACCCACCTCGCAAGTGCTCTAG | 2580 |
| Db | 2521 | TCGAGTCTTTGTAGATGGCAAACTCTGCACCTCTAAACCCACCTCGCAAGTGCTCTAG | 2580 |
| QY | 2581 | CCAGCGCTACATCTTTCTAGGCAAAATGCGCCCTTTGGGGCAGGTTATTTAGGAA | 2640 |
| Db | 2581 | CCAGCGCTACATCTTTCTAGGCAAAATGCGCCCTTTGGGGCAGGTTATTTAGGAA | 2640 |
| QY | 2641 | TCCTCATCAACCGCACCTGCAATCGGCTTCGCTGTCGATCGTGCTTTAAACCAAC | 2700 |
| Db | 2641 | TCCTCATCAACCGCACCTGCAATCGGCTTCGCTGTCGATCGTGCTTTAAACCAAC | 2700 |
| QY | 2701 | TTTTCATCTTTAAGCAATCCCATTTTAAATTAATTAATCTTATATTAATTAATATAT | 2760 |
| Db | 2701 | TTTTCATCTTTAAGCAATCCCATTTTAAATTAATTAATCTTATATTAATTAATATAT | 2760 |
| QY | 2761 | TATGCCCTCATTTTAAAGGAGAAATATGCGTAGGTCTTTGGTATTTGCTATGTGGGG | 2820 |
| Db | 2761 | TATGCCCTCATTTTAAAGGAGAAATATGCGTAGGTCTTTGGTATTTGCTATGTGGGG | 2820 |
| QY | 2821 | TTGTTTGGTGTGGGCGCAAGGGTATGAAACCCATCGGCTCAAAAAGTAGAAGCCAC | 2880 |
| Db | 2821 | TTGTTTGGTGTGGGCGCAAGGGTATGAAACCCATCGGCTCAAAAAGTAGAAGCCAC | 2880 |
| QY | 2881 | AGG 2883 | |
| Db | 2881 | AGG 2883 | |
| RESULT 2 | | | |
| AX356683 | | | |
| LOCUS | AX356683 | 2883 bp | DNA linear PAT 13-FEB-2002 |
| DEFINITION | Sequence 1 from Patent EP1176192. | | |
| ACCESSION | AX356683 | | |
| VERSION | AX356683.1 GI:18674020 | | |
| KEYWORDS | Helicobacter felis | | |
| SOURCE | Helicobacter felis | | |
| ORGANISM | Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Helicobacter. | | |
| REFERENCE | 1 | | |
| AUTHORS | Kusters, J.G. and Cattoli, G. | | |
| TITLE | Helicobacter felis vaccine | | |

| | |
|----------|---|
| JOURNAL | Patent: EP 1176192-A 1 30-JAN-2002; |
| FEATURES | Akzo Nobel N.V. (NL) |
| source | Location/Qualifiers |
| | 1..2883 |
| | /organism="Helicobacter felis" |
| | /mol_type="unassigned DNA" |
| | /db_xref="taxon:214" |
| | 206..886 |
| CDS | /note="unnamed protein product" |
| | /codon_start=1 |
| | /transl_table=11 |
| | /protein_id="CAD23255.1" |
| | /db_xref="GI:18674021" |
| | /translation="MKLTPKQEKFLLYVAGEVARKRKAEGKLKNOPEAIAYISAHIM |
| | DEARGKKTVAQLMEECMHFLKDEVMGPGVGNVPLDGEATFPDGTCLIVNWP |
| | DEHFKAGRVKFGCDKDIELNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKP |
| | KAYKRLDIPSGNTLIRIGAGQTRKVLPLGSGKKVIGNGLVNNIADERHKH |
| | AKSHGFYK" |
| | 897..2603 |
| | /note="unnamed protein product" |
| | /codon_start=1 |
| | /transl_table=11 |
| | /protein_id="CAD23256.1" |
| | /db_xref="GI:18674022" |
| | /translation="MKMKQEVNTYGYPTGDKVRLGDTDLMAEVBHDYTTYGEELKF |
| | GAGKTIREGMGQSNPDENLTLVITNAMIIDYGTIYKADIGIKNGIKHIGIKAGNKD |
| | MODGVSVMVVGVTGTEALAGEMGITAGGIDSHTHPLSPQOPTALANGVTTFMG |
| | GPVDTNATTTIPGKWNLRMLRAAEYSVMVYFLGKGNSSKKQOLVQVEAGAI |
| | LHEBWGTPSALIDHCLSVADYDVQVCIHTDVTNENAGYVDDTLNMGNGRAIHAYH |
| | AGGSHSPDVTIMAGELNLPSTPTPTINTVAHLDMLTCHHLDKRRIRDLP |
| | SKIRPISIAAEVDLHMGVIAMTSDSQAMGRAGEVIRPTQWADKNKKEFKLPED |
| | GKMDNFRIKRYISKYNINPALTHGVSEYISGVERGKIADLVNPNKPFKPKVIK |
| | GMVVFSPMGDSNASVPTPOPVYVREMFGHGKAKFDTSTTFVSKVAYENGVKELGL |
| | ERVLVVKNCRNITKKDFKENDTKAIVDPKTFEYFVDGKLTCKTSKPTSQVPLAQR |
| | YT" |

ORIGIN

| | | | | |
|-----------------------|--------------|---|------------|--------------|
| Query Match | 99.9% | Score 2880.6; | DB 6; | Length 2883; |
| Best Local Similarity | 100.0% | Pred. No. 0; | | |
| Matches 2883; | Conservative | 0; | Mismatches | 0; |
| | | | Indels | 0; |
| | | | Gaps | 0; |
| QY | 1 | RGGAGATTTTCCARCACTTCAAGCACATATTTGATCCTCTGTGTGGTGGTAAATTCR | 60 | |
| Db | 1 | RGGAGATTTTCCARCACTTCAAGCACATATTTGATCCTCTGTGTGGTGGTAAATTCR | 60 | |
| QY | 61 | ACTTGTTAATRCATATTAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTA | 120 | |
| Db | 61 | ACTTGTTAATRCATATTAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTA | 120 | |
| QY | 121 | TTACTTTATTAATAAAGTTAATAAAGTAAACGAAATTAGGACTATAATCCCATTCGCTT | 180 | |
| Db | 121 | TTACTTTATTAATAAAGTTAATAAAGTAAACGAAATTAGGACTATAATCCCATTCGCTT | 180 | |
| QY | 181 | TAAATTTAAACAAGGAGTAAAGTGAAGTAAAGTGAAGTGAAGTGAAGTGAAGTGAAG | 240 | |
| Db | 181 | TAAATTTAAACAAGGAGTAAAGTGAAGTAAAGTGAAGTGAAGTGAAGTGAAGTGAAG | 240 | |
| QY | 241 | GTTATATTTATCGGGCGAAGTGGCTAGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCA | 300 | |
| Db | 241 | GTTATATTTATCGGGCGAAGTGGCTAGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCA | 300 | |
| QY | 301 | ACCGAAGCCATTCGTTACATTTAGTCCCATATTTATGACGAGCGCGCGTGGAAAAA | 360 | |
| Db | 301 | ACCGAAGCCATTCGTTACATTTAGTCCCATATTTATGACGAGCGCGCGTGGAAAAA | 360 | |
| QY | 361 | AACGTTGCCAGCTTATGGAAGAGTGCATGCACTTTTGAATAAAGAGTGAAGTGAAGT | 420 | |
| Db | 361 | AACGTTGCCAGCTTATGGAAGAGTGCATGCACTTTTGAATAAAGAGTGAAGTGAAGT | 420 | |
| QY | 421 | CGGGGTGGGTAAATATGGTTCGCCGATCTAGGTAGAACCCACTTTTCTGATCGGTACG | 480 | |
| Db | 421 | CGGGGTGGGTAAATATGGTTCGCCGATCTAGGTAGAACCCACTTTTCTGATCGGTACG | 480 | |

QY 481 ACTTGTAACCTGTGAATTTGGCCCATCGAACAGATGAGCACTTCAAAAGCGGGCGAAGTGAA 540
Db 481 ACTTGTAACCTGTGAATTTGGCCCATCGAACAGATGAGCACTTCAAAAGCGGGCGAAGTGAA 540
QY 541 ATTTGGTTCGCGATAAAGACATCGAGCTCAATGCAGGCAAGAGTAACCGAACTTCAGGT 600
Db 541 ATTTGGTTCGCGATAAAGACATCGAGCTCAATGCAGGCAAGAGTAACCGAACTTCAGGT 600
QY 601 TACTAATGAAGGCGCTAAATTCCTTGCATGTGGGTAGCCATTTTCCACTTCTTTTGAAGCTAA 660
Db 601 TACTAATGAAGGCGCTAAATTCCTTGCATGTGGGTAGCCATTTTCCACTTCTTTTGAAGCTAA 660
QY 661 CAAAGGCACTAAATTCGATCGTGAAAGAACCGCTATGGCAAAAGCGCTAGATATTCCTCTCTGG 720
Db 661 CAAAGGCACTAAATTCGATCGTGAAAGAACCGCTATGGCAAAAGCGCTAGATATTCCTCTCTGG 720
QY 721 CAACACGCTACGCTATGGGCGAGGACAAACCCGCAAGTGCGAGTTCCTCTCTGGTGG 780
Db 721 CAACACGCTACGCTATGGGCGAGGACAAACCCGCAAGTGCGAGTTCCTCTCTGGTGG 780
QY 781 CAGTAAAAAGTGAATTTGGCATGAACCGGCTTTGTGAATAACATCGCGATGAACGCCATAA 840
Db 781 CAGTAAAAAGTGAATTTGGCATGAACCGGCTTTGTGAATAACATCGCGATGAACGCCATAA 840
QY 841 ACATAAAGCGCTTGAACAAGCGGAAATCTCAAGGATTTTCAAGTAAGGAGACTCCCATGA 900
Db 841 ACATAAAGCGCTTGAACAAGCGGAAATCTCAAGGATTTTCAAGTAAGGAGACTCCCATGA 900
QY 901 AAATGAAAAAACAAGTAATGTAAATACCTACGGACCCCAAGCGGATAAAGTGGCT 960
Db 901 AAATGAAAAAACAAGTAATGTAAATACCTACGGACCCCAAGCGGATAAAGTGGCT 960
QY 961 TAGGAGATACCGATCTTTGGCGAAGTAGAATGAACTATACCATCTATGGCGAAGAAC 1020
Db 961 TAGGAGATACCGATCTTTGGCGAAGTAGAATGAACTATACCATCTATGGCGAAGAAC 1020
QY 1021 TTAATTTTGGCGGGTAAACATATCCGTGAGGATGGGTGAGCAATAGCCCTGATG 1080
Db 1021 TTAATTTTGGCGGGTAAACATATCCGTGAGGATGGGTGAGCAATAGCCCTGATG 1080
QY 1081 AAAACACCTAGATTTAGTCACTACATACCGCATGATTTAGTCACTACCGGGATTTTACA 1140
Db 1081 AAAACACCTAGATTTAGTCACTACATACCGCATGATTTAGTCACTACCGGGATTTTACA 1140
QY 1141 AAGCCGACATTTGGATTTAAACCGGCAAAATCCATGGCATTTGGCAAGGAGGAAACAAG 1200
Db 1141 AAGCCGACATTTGGATTTAAACCGGCAAAATCCATGGCATTTGGCAAGGAGGAAACAAG 1200
QY 1201 ACATGCAAGATGGCGTAAGCCCTCATATGGTCTGGGTGTTGGGCACAGAGCACTAGCAG 1260
Db 1201 ACATGCAAGATGGCGTAAGCCCTCATATGGTCTGGGTGTTGGGCACAGAGCACTAGCAG 1260
QY 1261 GGAAGGTATGATTTACCGCTGGGGGAATCGATTCACACCCACTTCTCTTTTCTCCAC 1320
Db 1261 GGAAGGTATGATTTACCGCTGGGGGAATCGATTCACACCCACTTCTCTTTTCTCCAC 1320
QY 1321 AACAAATCCCTACCGCTAGCCAAATGGCGTTACACCAATGTTTGGAGGGCGGACAGGTC 1380
Db 1321 AACAAATCCCTACCGCTAGCCAAATGGCGTTACACCAATGTTTGGAGGGCGGACAGGTC 1380
QY 1381 CTGTAGATGGCAGAAATGGCACTACTATCACTCGGGCAAAATGGAATTCACCGCATGT 1440
Db 1381 CTGTAGATGGCAGAAATGGCACTACTATCACTCGGGCAAAATGGAATTCACCGCATGT 1440
QY 1441 TCGCGCAGCAGAAAGATATTTGATGAAATGGGCTTTTGGGCAAGGCAATAGCTCTTA 1500
Db 1441 TCGCGCAGCAGAAAGATATTTGATGAAATGGGCTTTTGGGCAAGGCAATAGCTCTTA 1500
QY 1501 GCNAAAAACNACTTTAGAACAAAGTAGAAGCGGGCGGATTTGGTTTTAAATTTGATGAAG 1560
Db 1501 GCNAAAAACNACTTTAGAACAAAGTAGAAGCGGGCGGATTTGGTTTTAAATTTGATGAAG 1560
QY 1561 ACTGGGGCACAACACCAAGTGCGATCGATCACTGCTTGGCGGTGGCGATGAATACGATG 1620

Db 1561 ACTGGGGCACAACCAAGTGGCATCGATCACTGCTTTAGCGTGGCAGATGAATACGATG 1620
QY 1621 TGCAAGTTTCTATCCACACCGATACAGTCAATGAGGCGAGGTTATGTAGATGACACCCCTAA 1680
Db 1621 TGCAAGTTTCTATCCACACCGATACAGTCAATGAGGCGAGGTTATGTAGATGACACCCCTAA 1680
QY 1681 ATGCAATGAACGGCGCGCCATCCATGCTTACCACTTGAAGGAGCGGGTGGAGSACACT 1740
Db 1681 ATGCAATGAACGGCGCGCCATCCATGCTTACCACTTGAAGGAGCGGGTGGAGSACACT 1740
QY 1741 CACTGATGTTATCACCATGCGAGGCGAGCTCAATATTTCTACCTCTCTCCACACCCCA 1800
Db 1741 CACTGATGTTATCACCATGCGAGGCGAGCTCAATATTTCTACCTCTCTCCACACCCCA 1800
QY 1801 CTATTCCTTATACCATTAATACGGTTGAGAACTTGTAGACATGCTCATGATGCGCAACC 1860
Db 1801 CTATTCCTTATACCATTAATACGGTTGAGAACTTGTAGACATGCTCATGATGCGCAACC 1860
QY 1861 ACCTAGACAAACGATCCGCGAGGATTTACAATTTTCTCAAAGCGTATCCGCCCCGGCT 1920
Db 1861 ACCTAGACAAACGATCCGCGAGGATTTACAATTTTCTCAAAGCGTATCCGCCCCGGCT 1920
QY 1921 CTATCGCGCTGAAGATGTGCTCATGATATGGGTGTATCGCATGACAAAGCTCGGATTT 1980
Db 1921 CTATCGCGCTGAAGATGTGCTCATGATATGGGTGTATCGCATGACAAAGCTCGGATTT 1980
QY 1981 CGCAAGCAATTTGGGCGGTGCGAGGCGAAGTGTCTCTCGAACTTTGGCAGACTCGGGATAA 2040
Db 1981 CGCAAGCAATTTGGGCGGTGCGAGGCGAAGTGTCTCTCGAACTTTGGCAGACTCGGGATAA 2040
QY 2041 ATAAAAAAGAAATTTGGTAAAGCTTCTCGAAGATGGCAAGATAACGATAAATTTCCGCAATTA 2100
Db 2041 ATAAAAAAGAAATTTGGTAAAGCTTCTCGAAGATGGCAAGATAACGATAAATTTCCGCAATTA 2100
QY 2101 AGCGCTACATCTCCAAATACATATCAACCCCGCTTTTGAACCCACGCGGTGAGCGAGTATA 2160
Db 2101 AGCGCTACATCTCCAAATACATATCAACCCCGCTTTTGAACCCACGCGGTGAGCGAGTATA 2160
QY 2161 TCGGCTCTGTGGAGAGGGCAAGATCGCGACTTTGGTGTGTGGAATCTCGCTTTTTTTG 2220
Db 2161 TCGGCTCTGTGGAGAGGGCAAGATCGCGACTTTGGTGTGTGGAATCTCGCTTTTTTTG 2220
QY 2221 GCGTAAAAACCAAAATCGTATCAAGCGGTATGTGGTCTTCTCTGAAATGGCGGATTT 2280
Db 2221 GCGTAAAAACCAAAATCGTATCAAGCGGTATGTGGTCTTCTCTGAAATGGCGGATTT 2280
QY 2281 CTAAACGCTCTGTGCCCACTCCCAACCGGTTTATTACCGGAAATGTTTGGGCATCAG 2340
Db 2281 CTAAACGCTCTGTGCCCACTCCCAACCGGTTTATTACCGGAAATGTTTGGGCATCAG 2340
QY 2341 GCAAGGCGAAATTTGACACCGAGTCACTTTTGTTCGCGGTCAAAACCTGCGTAAACATCA 2400
Db 2341 GCAAGGCGAAATTTGACACCGAGTCACTTTTGTTCGCGGTCAAAACCTGCGTAAACATCA 2400
QY 2401 TGAAGAAAAAGCTGGGCTTTAGAGCGCCAAAGTTCTACCGGTCAAAACCTGCGTAAACATCA 2460
Db 2401 TGAAGAAAAAGCTGGGCTTTAGAGCGCCAAAGTTCTACCGGTCAAAACCTGCGTAAACATCA 2460
QY 2461 CCAAGAAAGACTTCAAGTTTCAACCGCAAAACCGGCAAAATCACCGTCGATCCGAAAACT 2520
Db 2461 CCAAGAAAGACTTCAAGTTTCAACCGCAAAACCGGCAAAATCACCGTCGATCCGAAAACT 2520
QY 2521 TCAGAGTCTTTGTAGATGGGCAAACTCTGCACTCTAAACCCACCTCGCAAGTGCCTCTAG 2580
Db 2521 TCAGAGTCTTTGTAGATGGGCAAACTCTGCACTCTAAACCCACCTCGCAAGTGCCTCTAG 2580
QY 2581 CCCAGCGCTACATTTCTTCTAGGACAAATGCCCCCTTTGGGGCAGGTTATTTTAGGAA 2640
Db 2581 CCCAGCGCTACATTTCTTCTAGGACAAATGCCCCCTTTGGGGCAGGTTATTTTAGGAA 2640
QY 2641 TCTTCATCAACACGCACTCGCAATCGGTCTTTCGCTGTGCGATCGTGTCTGCTTTTAAACCAAC 2700

Db 2641 TCTTCATCAACGCACTGCAATCGGTCTTGGGTGCGATCGTGCCTTTAAACAAC 2700
Qy 2701 TTTTCATCTTTAAAGCAATGCCCAATTTTAAATTAATTAATTTAAATTTTAAATTAATTAAT 2760
Db 2701 TTTTCATCTTTAAAGCAATGCCCAATTTTAAATTAATTAATTTTAAATTTTAAATTAATTAAT 2760
Qy 2761 TATGCCCTCATTTTAAAGGAGAATATGCGTAGGTCTTTGGTATTTGCTATGCTGCGGG 2820
Db 2761 TATGCCCTCATTTTAAAGGAGAATATGCGTAGGTCTTTGGTATTTGCTATGCTGCGGG 2820
Qy 2821 TTGTTTGGTCTGGCGCAAGAGGTATTGAAACCCATCGCTCAAAAAAGGTAGAAAGCCAC 2880
Db 2821 TTGTTTGGTCTGGCGCAAGAGGTATTGAAACCCATCGCTCAAAAAAGGTAGAAAGCCAC 2880
Qy 2881 AGG 2883
Db 2881 AGG 2883

RESULT 3
BD185306
LOCUS BD185306 2452 bp DNA linear PAT 17-JUN-2003
DEFINITION Helicobacter felis vaccine.
ACCESSION BD185306
VERSION BD185306.1 GI:31877506
KEYWORDS JP 2002355054-A/5.
SOURCE Helicobacter felis
ORGANISM Helicobacter felis
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Helicobacteraceae; Helicobacter.
REFERENCE 1 (bases 1 to 2452)
AUTHORS Kusters,J.G. and Cattoli,G.
TITLE Helicobacter felis vaccine
JOURNAL Patent: JP 2002355054-A 5 10-DEC-2002;
AKZO NOBEL NV
COMMENT OS Helicobacter felis
PN JP 2002355054-A/5
PD 10-DEC-2002
PF 16-JUL-2001 JP 2001214711
PR 17-JUL-2000 EP 00202565.8
PI JOHANNES GERARDUS KUSTERS,GIOVANNI CATTOLI
PC C12N15/09,A61K38/00,A61K39/106,A61K39/118,A61K39/12,A61K39/175, PC
A61K39/23
PC A61K39/235,A61K39/39,A61K39/395,A61P1/04,A61P31/04,C12N1/15,
C12N1/19,
PC C12N1/21,C12N5/10,C12N9/80,C12Q1/68,G01N33/15,G01N33/50,G01N33/ PC
53,
PC G01N33/53,G01N33/566,G01N33/569/(C12N9/80,C12R1/01),(C12Q1/68, PC
C12R1/01),
PC C12N15/00,C12N5/00,A61K37/02
CC Helicobacter felis vaccine
FH Key Location/Qualifiers
FT CDS (48)..(728)
FT CDS (739)..(2445).

FEATURES
source
1..2452
/organism="Helicobacter felis"
/mol_type="genomic DNA"
/db_xref="taxon:214"

ORIGIN
Query Match 77.3%; Score 2228; DB 6; Length 2452;
Best Local Similarity 94.3%; Pred. No. 0;
Matches 2312; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
Qy 159 AGGACTATAATCCATTCGCTTTAAATTTAAACAAGGAGTAAATAGGTGAAACTCAC 218
Db 1 AGGACTATAATCCATTCGCTTTAAATTTAAACAAGGAGTAAATAGGTGAAACTCAC 60
Qy 219 CCAAAGACNAGAAAAGTCTTGTATATTATGCGGCGAAGTGGCTAGAAAGCGCAAG 278

Db 61 CCAAAGACNAGAAAAGTCTTGTATATTATGCGGCGAAGTGGCTAGAAAGCGCAAG 120
Qy 279 CAGAGGGCTTAAAGCTCAACCAACCCGAAGCCATTGCTTTACATTAGTGCCCATATTATGG 338
Db 121 CAGAGGGCTTAAAGCTCAACCAACCCGAAGCCATTGCTTTACATTAGTGCCCATATTATGG 180
Qy 339 ACGAAGCGCGCTGGGAAAAAAACCGTGTGCCAGCTTATGGAAGAGTGCATGCATTTT 398
Db 181 ACGAAGCGCGCTGGGAAAAAAACCGTGTGGGAACTTATATGGAAGAGTGTATGCACTTTT 240
Qy 399 TGA AAAAGATGAAGTAAATGCGCGGGTGGGTAAATATGTTTCCCGATCTAGTGTAGAAAG 458
Db 241 TGA AAAAGATGAAGTAAATGCGCGGGTGGGGAATATGTTTCCCTGATTTGGGCGTGAAG 300
Qy 459 CCACCTTTCTGTATGTTACGAAACTTGTAACTGTGAAATTTGGCCCATCGAACCCAGATGAGC 518
Db 301 CCACCTTTCCCGATGGCACCAACTCTGTAACCTGTAATTTGGCCCATCGAACCTGATGAC 360
Qy 519 ACTTCAAAAGCGGCGAAGTGAAATTTGTTGCGATAAAGACATCGAGCTCAATGCAAGCA 578
Db 361 ACTTTAAGCGGGTGAAGTGAAATTTGGCTGTGATAAAGACATTGAACTCAACGCGAGTA 420
Qy 579 AAGAAGTAAACCGAACTTGAGGTACTATGAAGGCGCTAAATCTTGCATGTGGGTAGCC 638
Db 421 AGGAAGTTTACCGAACTAGAACTTAAACGAAGGACCTAAATCTTGCATGTGGGTAGCC 480
Qy 639 ATTTTCCACTTCTTTGAAGCTAAACGACACTAAATTCGATCGTCAAAAAGCTTATGGCA 698
Db 481 ATTTTCCACTTCTTTGAAGCCAAACAAAGCATTTGAAATTCGATCGGAAAAAAGCTTATGGCA 540
Qy 699 AACGCTAGATATTTCCCTCTGCGCAACACGCTACGCAATTTGGGCGAGGACAAACCCGCAAG 758
Db 541 AACGCTAGATATTTCCCTCTGCGCAACACACTACGCAATTTGGGCGAGGACAAACCCGTAAG 600
Qy 759 TGCAGTTGATTCCTTTGGTGGCGAGTAAAAAGTGAATTTGGCATGAACCGGCTTTGGAATA 818
Db 601 TGCAGTTAAATCCCTCTTGGCGGTAGTAAAAAGTGAATTTGGCATGAACCGGCTTTGGAATA 660
Qy 819 ACATCGCGGATGAAGCCGCAATGAAGCGCTTGACAAGCGCAAAATCTCAGCGATTTA 878
Db 661 ATATTTGGCGAGCAAGCGCCATAACCAAAAGCGCTAGACAAAGCAAAATCTCAGCGATTTA 720
Qy 879 TCAAGTAGGAGACTCCCATGAAAATGAAAAAACAAGAAATATGTAATACCTACCGACCC 938
Db 721 TCAAGTAGGAGACTCCCATGAAAATGAAAAAACAAGAGTATGTAATACCTACCGACCC 780
Qy 939 ACCAAAGCGGATAAAGTGCCTTAGGAGATACCGATCTTTTGGGCGAGAGTAGAACATGAC 998
Db 781 ACCACAGCGGATAAAGTGCCTTAGGAGATACCGATCTTTTGGGCGAGAGTAGAACATGAC 840
Qy 999 TATACCACTATGCGGAAGAACTTAAATTTGGCGGGGTAAACATCTCCGTGAGGATAG 1058
Db 841 TATACCACTATGCGGAAGAACTCAAAATTCGCTGCAAGTAAACATCTCCGTGAGGATAG 900
Qy 1059 GGTCAAGCAATAGCCCTGATGAAAACACCCCTAGATTTAGTTCATCACTAACCCGATGAT 1118
Db 901 GGTCAAGCAATAGCCGAGTAAACACCTTTAGATTTAGTATCACCACCGCATGAT 960
Qy 1119 ATCGACTACACCGGATTTTACAAAGCCGACATTTGGGATTTAAACCGCAAAATTCATGCG 1178
Db 961 ATTGACTACACCGGATTTTACAAAGCCGACATTTGGCATTTAAATGCGCAAAATTCATGCG 1020
Qy 1179 ATTGCAAGGCGAAGAAACAGGACATGCAAGATGCGGTAAAGCCCTCATATGTCGTGGGT 1238
Db 1021 ATTGCAAGGCGAAGAAACAGGACATGCAAGATGCGGTAAAGCCCTCATATGTCGTGGGT 1080
Qy 1239 GTGGGCAAGAGACACTAGCAGGGGAGGTATGATTTATACCGTGGGGGATCCGATTCA 1298
Db 1081 GTGGGCAAGAGACACTAGCAGGGGAGGTATGATTTATACCGTGGGGGATCCGATTCA 1140
Qy 1299 CACACCCACTTCTTCTTCCCAACAATTCCTACCGCTCTAGCCCAATGGCGCTTACCAAC 1358

Db 1141 CACACCCACTTCTCTCTCCACAACAAATTCCTTACCGCTCTAGCCAAATGGCGTTACACA 1200
Qy 1359 ATGTTTGGAGCGGCACAGGTCTGTAGATGGCAGCAATCGCACTACTATCACTCCGGGC 1418
Db 1201 ATGTTTGGCGGTGGCACAGCCCGGTAGATGGCAGCAATCGCACTACTCCGGGC 1260
Qy 1419 AAATGGAATCTGCACCGCATGTTGCGCGAGCAGCAAGAGTATTTCTATGAATGTGGCTTT 1478
Db 1261 AAATGGAATCTGCACCGCATGTTGCGCGAGCAGCAAGAGTATTTCTATGAATGTGGCTTT 1320
Qy 1479 TTGGCAAGGCAATAGCTCTAGCAAAAACAACTTGTAGACAAGTGAAGCGGCGCG 1538
Db 1321 TTGGCAAGGCAATAGCTCTAGTAAAAACAACTTGTAGACAAGTGAAGCGGCGCG 1380
Qy 1539 ATTGGTTTAAATTTGCATGAAGACTGGGCAACAACCAAGTGCAGTCGATCACTGCTTG 1598
Db 1381 ATTGGTTTAAATTTGCATGAAGACTGGGCAACAACCTTCAAGTGCAGTCGATCACTGCTTG 1440
Qy 1599 AGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATACAGTCAATGAGGCA 1658
Db 1441 AGCGTAGCAGATGAATACGATGTGCAAGTTTGTATACACACCGATACAGTCAATGAGGCA 1500
Qy 1659 GGTATGTATGATGACACCCCTAAATGCAATGAACGGCGCGCATCCATGCCTACCAATT 1718
Db 1501 GGTATGTATGATGACACCCCTAAATGCAATGAACGGCGCGCATCCATGCCTACCAATT 1560
Qy 1719 GAGGAGCGGTGGAGGACACTCACCTGATGTTATCACCATGCGAGCGAGCTCAATATT 1778
Db 1561 GAGGAGCGGTGGAGGACACTCACCTGATGTTATCACCATGCGAGCGAGCTCAATATT 1620
Qy 1779 CTACCTCTCCACCAACCCCACTATTCCTATACCAATTAATACGTTGCGAACACTTA 1838
Db 1621 CTACCTCTCCACCAACCCCACTATTCCTATACCAATTAATACGTTGCGAACACTTA 1680
Qy 1839 GACATGCTCATGACATGCCACCACTAGACAACCGCATCGCGAGGATTTACAATTTCT 1898
Db 1681 GACATGCTTATGACCTGCCACCACTAGATAAACCGCATCGCGAGGATTTCCAATTTCT 1740
Qy 1899 CAAAGCGGTATCGCCCGGCTCTATCGCGGCTGAAGATGTGCTCCATGATATGGGTGTG 1958
Db 1741 CAAAGCGGTATCGCCCGGCTCTATCGCGGCTGAAGATGTGCTCCATGATATCGGTGTG 1800
Qy 1959 ATCGCGATGACAAGCTCGGATTCGCAAGCAATGGGCGGTGCGAGCGAAGTATTCCTCGA 2018
Db 1801 ATCGCGATGACAAGTTCCGATTCGCAAGCAATGGGCGGTGCGGGAAGTATTCCTCGA 1860
Qy 2019 ACTTGGCAGATCGCGATGAAGATAAAGAAATTTGGTAAGCTTCTGGAAGTGGCAAA 2078
Db 1861 ACTTGGCAAACTGCAGACAAGATAAAGAAATTTGGTAAGCTTCTGGAAGTGGTGCA 1920
Qy 2079 GATAACGATAATTTCCGCAATTAAGCGCTACATCTCCAAATACACTATCAACCCGCTTTCG 2138
Db 1921 GATAATGACAATTTCCGCAATTAAGCGCTATATCTCCAAATACACTATTAATCCCGCTTTCG 1980
Qy 2139 ACCACGGGTGAGCGAGTATATCGGCTCTGTGGAAAGAGGCGCAAGATCGCGCACTTGGTG 2198
Db 1981 ACCATGGCGTGAGCGAGTATATCGGCTCTGTGGAAAGAGGCGCAAGATCGCGCACTTGGTG 2040
Qy 2199 GTGTGAATCTCGCTTTTGGGTGTAACCCAAATCTGTATCAAAAGCGGTATGGTG 2258
Db 2041 GTGTGAATCTCGCTTTTGGGTGTAACCCAAATCTGTATCAAAAGCGGTATGGTG 2100
Qy 2259 GTCTTCTCTGAATGGCGAATTTCAACGGTCTGTGCCACTCCCAACCGGTTTATTAC 2318
Db 2101 GTGTCTCTGAATGGCGAATTTCAATGGCTGTGTGCCCACTCCTCAGCGCGTTTATTAC 2160
Qy 2319 CGCGAAATGTTTGGGCATACCGCAAGCGGAAATTTGACACCAAGCATCACTTTTGTTC 2378
Db 2161 CGCGAAATGTTTGGGCATACCGCAAGCGGAAATTTGACACCAAGCATCACTTTTGTTC 2220
Qy 2379 AAAGTCGCTATGAAATGCGGTGAAGAAAGCTGGGCTTAGAGCGCAAGTTCTACCG 2438
Db 2221 AAAGTCGCTATGAAATGCGGTGAAGAAAGAACTAGGTTTAGAGCGCAAGGTGCTCCCC 2280

Qy 2439 GTCAAAAACCTGCGTAACATCAACAAGAAAGACTTCAAGTTCAACGACAAAACGGCAAAA 2498
Db 2281 GTGAAAAACCTGCGTAACATCAACAAGAAAGACTTCAAGTTCAACGACAAAACGGCAAAA 2340
Qy 2499 ATCACCGTCGATCGGAAAACCTTCGAGGTCTTTGTAGATGGCAAACTCTGCACCTCTAAA 2558
Db 2341 ATCACCGTCGATCGGAAAACCTTCGAGGTCTTTGTAGATGGCAAACTCTGCACCTCTAAA 2400
Qy 2559 CCCACCTCGCAAGTGCCTCTAGCCCGAGCTGTACACTTTTCTTAGGCACAAT 2610
Db 2401 CCCACCTCGCAAGTGCCTCTAGCCCGAGCTGTACACTTTTCTTAGGCACAAT 2452

RESULT 4
AX356695
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
CDS
ORIGIN

AX356695
Sequence 13 from Patent EP1176192.
AX356695
AX356695.1 GI:18674032
Helicobacter felis
Helicobacter felis
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Helicobacteraceae; Helicobacter.
1
Kusters,J.G. and Cattoli,G.
Helicobacter felis vaccine
Patent: EP 1176192-A 13 30-JAN-2002;
Akzo Nobel N.V. (NL)
Location/Qualifiers
1..2452
/organism="Helicobacter felis"
/mol_type="unassigned DNA"
/db_xref="taxon:214"
48..728
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD23263.1"
/db_xref="GI:18674033"
/translation="MKLTPKEQKFLYYAGVARKKAEGLKLNQPEAIAYISAHIM
DEARRKKTVAELMECEMFLKKDEVNPGVMVPLDGEATFPDKTKLVNWNVIEP
DEHFKAEGVFQCDKDIELNAGKEVTELVNEGPKSLHVGSHFPEFANKALKFRE
KATGKRLDIPSGNTRLRIGAGQTRKVLQILPGGSKKVIQMGNLVNNIADERHKHKALDK
AKSHGPIK"
739..2445
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD23264.1"
/db_xref="GI:18674034"
/translation="MKMKQKQYVNTYPTGDKVRLGDTDLWAEVEHDYTYTGEELKF
GAKTIREGMGNSPDENTLDLVTNAMIIDYGIYKADIGIKNGKIHGIGKGNKD
MQQGVPMVVGVTGEALAGEMIITAGGIDSHFSLSPQFPQFALANGVTMTFGGCT
GPVDGNATTITPGKWNLRMLRAAEVSMNVGFLGKNSSSKKQLVQVBEAGAIGFK
LHEDMGTPSAIDHCLSVADBYDVQVCIHTDTVNEAGYVDDTLNMGRAIHAHIEG
AGGHSPPDVIIMAGEVNLPSSTPTTPIYINTVAEHLDMLTCHLDRKREDLQFS
QSRIRPGSIAEDVLHDIGVIAMTSSDSQAMGRAGEVIRPTWQADKNKKGFKLUPD
GANDNFRIKRYISKYINPALTHVSEYIGSVEEGEADIADVNNWPAFVGKPKTVIK
GMVVFSEMGDSNASVTPQPVYVREMFHGHGKAFDTSITFVSKVAYENGVKELGL
ERKVLPVKNCRNITKKDKFNDKTAKITVDPKTFEVFDGKLTCKTSKPTSEVPLAQRYT
PF"

ORIGIN
Query Match 77.3%; Score 2228; DB 6; Length 2452;
Best Local Similarity 94.3%; Pred. No. 0;
Matches 2312; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
Qy 159 AGGACTATAATCCCATTCGCTTTAAATTTAAACAAGAGTAAATAGGTGAAACTCACAC 218
Db 1 AGGACTATAATCCCATTCGCTTTAAATTTAAACAAGAGTAAATAGGTGAAACTCACAC 60

Qy 219 CCAAAGACGAAGAAAGTTCTTGTATATATATGCGGCGAAGTGCGCTAGAAAGCGCAAG 278
Db 61 CCAAAGACGAAGAAAGTTCTTGTATATATATGCGGCGAAGTGCGCTAGAAAGCGCAAG 120
Qy 279 CAGAGGCTTAAAGCTCAACCAACCGAAGCCATTGCTTACATTAGTGCCCATATTATGG 338
Db 121 CAGAGGCTTAAAGCTCAACCAACCGAAGCCATTGCGCTACATTAGTGCCCATATTATGG 180
Qy 339 ACGAAGCGCGCTGGGAAAAAACCCTTGCAGCTTATGGAAGAGTGCAATGCACTTTT 398
Db 181 ACGAAGCGCGCTGGGAAAAAACCCTTGCAGCTTATGGAAGAGTGATGCACTTTT 240
Qy 399 TGAAGAAAGATGAAGTAAATGCCCGGGTGGGTAAATATGGTTCGCACTAGGTGTAGAG 458
Db 241 TGAAGAAAGACGAGGTGATGCCCGGGTGGGGAATATGGTCCCTGATTTGGCGGTGGAAG 300
Qy 459 CCACCTTTCCTGATGGTACGAACCTTGTAACTGTGAATTTGGCCCATCGAACCATGAGC 518
Db 301 CCACCTTTCGCGATGGCACCAACTCGTAATCTGTGAATTTGGCCCATCGAACCTGATGAAC 360
Qy 519 ACTTCAAAGCGGCGAAGTGAATTTGCTGCGATAAAGACATCGAGCTCAATGCAGCA 578
Db 361 ACTTTAAGCGGCTGAAGTGAATTTGCTGTGATAAAGACATTTGAACCTCAACGAGTA 420
Qy 579 AAGAAGTAAACCGAATTTGAGGTTACTAATGAAGGCGCTAAATCCTTGCATGTGGGTAGCC 638
Db 421 AGGAAGTTACCGAATCTAGAGTTACTAAACGAGGACCTAAATCCTTGCATGTGGGTAGCC 480
Qy 639 ATTTCCACTTCTTTGAAGCTTAAACGAGCACTAAATTCGATCGTGAAAAAGCCCTATGGCA 698
Db 481 ATTTCCACTTCTTTGAAGCCAAACAAAGCAATTTGAATTCGATCGGGAAGAAAGCCCTATGGCA 540
Qy 699 AAGCCTTAGATATCCCTCTGGCAACACGCTACGATTTGGGCGAGGCAACCCGCAAG 758
Db 541 AAGCCTTAGATATCCCTCTGGCAACACACTAOCATTTGGGCGAGGCAAAACCCGTAAG 600
Qy 759 TGCAGTTGATTCCTCTTGGTGGCAGTAAAAAGTGAATTTGGCATGAAACCGGCTTGTGAATA 818
Db 601 TGCAGTTAATCCCTCTTGGCGGTAGTAAAAAGTGAATTTGGCATGAAACCGGCTTGTGAATA 660
Qy 819 ACATCGCGGATGAACGCCATAAATGAAGCGCTTGAACAGGCGAAATCTCACGGATTTA 878
Db 661 ATATTGGCGAAGCAAGCCATAAACACAAAGCGCTAGACAAAGCAAAATCTCACGGATTTA 720
Qy 879 TCAAGTAAGAGACTCCATGAAAAATGAAAAACAAAGATATGTAATATCCTACGAGACC 938
Db 721 TCAAGTAAGAGACTCCCATGAAAAATGAAAAACAAAGATATGTAATATCCTACGAGACC 780
Qy 939 ACCAAAGGCGATAAAGTGCCTTAGGAGATACCGATCTTTGGGCGAAGTAGAACATGAC 998
Db 781 ACCAAGGCGATAAAGTGCCTTAGGAGATACCGATCTTTGGGCGAAGTAGAACATGAC 840
Qy 999 TATACCACTTATGGCGAAGAACTTAAATTTGGCGCGGTAAAACTATCCGTGAGGGTATG 1058
Db 841 TATACCACTTATGGCGAAGAACTCAATTCGGTGCAGGTAAAACTATCCGTGAGGGTATG 900
Qy 1059 GGTACAGCAATAGCCCTGATGAAAAACACCCCTAGATTATGTCATCACTAACCGCATGATT 1118
Db 901 GGTACAGCAATAGCCCGATGAAAAACACCTTAGATTATGATGATCACCACCGCATGATT 960
Qy 1119 ATCGACTACACCGGATTTACAAAGCCGACATTTGGGATTTAAAAACGGCAAAATCCATGGC 1178
Db 961 ATTGACTACACCGGATTTACAAAGCCGACATTTGGCATTTAAAAATGGCAAAATCCATGGC 1020
Qy 1179 ATTGGCAAGCAGCAAAACAAAGGACATGCAAGATGGCGTAAAGCCCTCATATGGTCTGGGT 1238
Db 1021 ATTGGCAAGCAGCAAAACAAAGGACATGCAAGATGGCGTAAAGCCCTCATATGGTCTGGGT 1080
Qy 1239 GTGGGACAGAAGCACTAGCAGGGGAAGGTATGATTTATCCGCTGGGGAAATCGATTCA 1298
Db 1081 GTGGGACAGAAGCACTAGCAGGGGAAGGTATGATTTATCCGCTGGGGGATCGATTCA 1140
Qy 1299 CACACCCACTTCTTCTCCAAACAAATTCCTTACCCTCTAGCCATGGCGGTTACAAAC 1358

Db 1141 CACACCCACTTCTCTCTCCAAACAAATTCCTTACCGCTCTAGCCCAATGGCGTTTACAACA 1200
Qy 1359 ATGTTTGAGGGCGGCACAGGTCCTGTAGATGCAACGAATGCGACTACTATCACTCTCGGGC 1418
Db 1201 ATGTTTGCGGTGGGCACAGGCCCGGTAGATGCGACGAATGCGACTACCATCACTCTCGGGC 1260
Qy 1419 AAATGGAACTTGCACCGCATGTTGCGCGCAGCAGAGAGTATTTCTATGAATGTGGGCTTT 1478
Db 1261 AAATGGAACTTGCACCGCATGTTGCGCGCAGCAGAGAGTATTTCTATGAATGTGGGCTTT 1320
Qy 1479 TTGGGCAAGGCAATAGCTCTAGCAAAAAAACAATTTGTAGAACAAAGTAGAGCGGGCGG 1538
Db 1321 TTGGGCAAGGCAATAGCTCTAGTAAAAAACAATTTGTAGAACAAAGTAGAGCGGGCGG 1380
Qy 1539 ATTGGTTTAAATTCGATGAAGACTGGGGCAACAACCAAGTGCATCGATCACTGCTTG 1598
Db 1381 ATTGGTTTAAATTCGATGAAGACTGGGGCAACAATCCTAAGTGCATCGATCACTGCTTG 1440
Qy 1599 AGCGTGGCAGATGAATAGCATGTGCAAGTTTGTATCCACACCGATACAGTCAATGAGGCA 1658
Db 1441 AGCGTAGCAGATGAATAGCATGTGCAAGTTTGTATACACCGATACAGTCAATGAGGCA 1500
Qy 1659 GGTATGTAGATGACACCCCTAAATGCAATGAACGGGCGCCATCCATGCTTACCACTT 1718
Db 1501 GGTATGTAGATGACACCCCTAAATGCAATGAACGGGCGCCATCCATGCTTACCACTT 1560
Qy 1719 GAGGAGCGGTGGAGGACACTCACTGATGTTATCACCATGGCAGGCGAGTCAATATT 1778
Db 1561 GAGGAGCGGTGGAGGACACTCACTGATGTTATCACCATGGCAGGCGAGTCAATATT 1620
Qy 1779 CTACCTCTCCACACCCCACTATTCCTATACCATTAATACCGTGTGAGAACTT 1838
Db 1621 CTACCTCTCCACACCCCACTATTCCTATACCATTAATACCGTGTGAGAACTT 1680
Qy 1839 GACATGCTCATGACATGCGCACCCCTAGACAAACGATCCGAGGAGATTTACAATTTCT 1898
Db 1681 GACATGCTCATGACATGCGCACCCCTAGATAAACGATCCGAGGAGTCTCCAATTTCT 1740
Qy 1899 CAAAGCGTATCCGCGCGGCTCTATCGGGCTGAAGATGTGCTCATGATATGGGTG 1958
Db 1741 CAAAGCGTATCCGCGCGGCTCTATCGCGCTGAAGATGTGCTCATGATATCGGTG 1800
Qy 1959 ATCGGATGACAAAGCTCGGATTCGCAACAAATGGGCGGTGAGGCGAAGTCTCTCGA 2018
Db 1801 ATCGGATGACAAAGTTCGATTCGCAACAAATGGGCGGCTGGGGAAGTATCTCTAGA 1860
Qy 2019 ACTTGGCAGACTGCGGATAAGAAATAAAAAAGAAATTTGGTAAGCTTCTCGAAGATGCAAA 2078
Db 1861 ACTTGGCAAACTGCGAGACAGAAATAAAAAAGAAATTTGGTAAGCTTCTCGAAGATGCA 1920
Qy 2079 GATAACGATAATTTCCGCAATTAAGCGCTACATCTCAAAATAACATATCAACCCCGCTTG 2138
Db 1921 GATAATGACAACTTCGCGATCAACCGCTATATCTCAAAATAACCAATTAATCCCGCTTG 1980
Qy 2139 ACCACGCGGTGAGCGAGTATATCGGCTCTGTGGAAGGGGCAAGATCCCGACTTGGT 2198
Db 1981 ACCATGCGGTGAGCGAGTATATCGGCTCTGTGGAAGGGGCAAGATCCCGACTTGGT 2040
Qy 2199 GTGTGGAATCTGCTCTTTTGGCGTAAAAACCAAAATTCGTGATCAAAAGCGGTATGGTG 2258
Db 2041 GTGTGGAATCTGCTCTTTTGGCGTAAAAACCAAAATTCGTGATCAAAAGCGGTATGGTG 2100
Qy 2259 GTCTTCTGAAATGGGCGATTCTAACCGCTCTGTGCCACTCCCCCAACCGGTTTATTAC 2318
Db 2101 GTGTTCTCTGAAATGGGCGATTCTAAATGCGTCTGTGCCACTCTCAGCGGTTTATTAC 2160
Qy 2319 CGGAAATGTTTGGGCGATCAGCGCAAGCGGAAATTTGACACGAGCATCACTTTGTTTC 2378
Db 2161 CGGAAATGTTTGGGCGATCAGCGCAAGCGGAAATTTGACACGAGCATCACTTTGTTTC 2220
Qy 2379 AAAGTGCCTATGAAATGGCGTGAAGAAAGCTGGGCTTAGAGCGGCAAGTTCCTACCG 2438

| | | | |
|-----------------------|--|--|----------------------------|
| Db | 2221 | AAAGTCGCCTATGAAAATGGTGTGAAGAAAACCTAGGTTTAGAGCGCAAGGTGCTCCCC | 2280 |
| Qy | 2439 | GTCAAAAACCTGCGGTAAACATCACCAAGAAAGACTTCAAGTTCAACGACAAAACGGCAAAA | 2498 |
| Db | 2281 | GTGAAAACCTGCGGTAAACATCACCAAGAAAGACTTCAAGTTCAACGACAAAACGGCAAAA | 2340 |
| Qy | 2499 | ATCACCGTGCATCCGAAAACCTTGGAGGCTTTGTAGATGGCAAACTCTGCACCTCTAAA | 2558 |
| Db | 2341 | ATCACCGTGCATCCGAAAACCTTGGAGGCTTTGTAGATGGCAAACTCTGCACCTCTAAA | 2400 |
| Qy | 2559 | CCACCTGCAAGTGCCTCTAGCCAGCGCTACACTTCTCTAGGCACAAT | 2610 |
| Db | 2401 | CCACCTGCAAGTGCCTCTAGCCAGCGCTACACTTCTCTAGGCACAAT | 2452 |
| RESULT 5 | | | |
| LOCUS | BD185303 | 2405 bp | DNA linear PAT 17-JUN-2003 |
| DEFINITION | Helicobacter felis vaccine. | | |
| ACCESSION | BD185303 | | |
| VERSION | BD185303.1 | GI:31877503 | |
| KEYWORDS | JP 2002355054-A/2. | | |
| SOURCE | Helicobacter felis | | |
| ORGANISM | Helicobacter felis | | |
| REFERENCE | 1 (bases 1 to 2405) | | |
| AUTHORS | Kusters, J. G. and Cattoli, G. | | |
| TITLE | Helicobacter felis vaccine | | |
| JOURNAL | Patent: JP 2002355054-A 2 10-DEC-2002; | | |
| COMMENT | AKZO NOBEL NV | | |
| | OS Helicobacter felis | | |
| | PN JP 2002355054-A/2 | | |
| | PD 10-DEC-2002 | | |
| | PF 16-JUN-2001 JP 2001214711 | | |
| | PR 17-JUL-2000 EP 00202565.8 | | |
| | PI JOHANNES GERARDUS KUSTERS, GIOVANNI CATTOLI | | |
| | PC | | |
| | C12N15/09, A61K38/00, A61K39/106, A61K39/118, A61K39/12, A61K39/175, PC | | |
| | A61K39/23 | | |
| | PC A61K39/235, A61K39/39, A61K39/395, A61P1/04, A61P31/04, C12N1/15, | | |
| | PC C12N1/19, | | |
| | PC | | |
| | C12N1/21, C12N5/10, C12N9/80, C12Q1/68, G01N33/15, G01N33/50, G01N33/ PC | | |
| | 53, | | |
| | PC | | |
| | G01N33/53, G01N33/566, G01N33/569// (C12N9/80, C12R1.01), (C12Q1/68, PC | | |
| | C12R1.01), | | |
| | PC C12N15/00, C12N5/00, A61K37/02 | | |
| | CC Helicobacter felis vaccine | | |
| | PH Key Location/Qualifiers | | |
| | FT CDS (1)..(681) | | |
| | FT CDS (692)..(2398). | | |
| FEATURES | Location/Qualifiers | | |
| source | 1..2405 | | |
| | /organism="Helicobacter felis" | | |
| | /mol_type="genomic DNA" | | |
| | /db_xref="taxon:214" | | |
| ORIGIN | | | |
| Query Match | 74.9%; | Score 2158.6; | DB 6; Length 2405; |
| Best Local Similarity | 93.6%; | Pred. No. 0; | |
| Matches 2251; | Conservative 0; | Mismatches 154; | Indels 0; Gaps 0; |
| Qy | 206 | GTGAAACTCACACCCAAAGAGCAAGAAAGTCTTGTATATATATGCGGCGAAGTGGCT | 265 |
| Db | 1 | GTGAAACTCACACCCAAAGAGCAAGAAAGTCTTGTATATATATGCGGCGAAGTGGCT | 60 |
| Qy | 266 | AGAAAGCGCAAGCAGAGGCTTAAAGCTCAACCAACCCGAAGCCATTGCTTACATTAGT | 325 |
| Db | 61 | AGAAAGCGCAAGCAGAGGCTTAAAGCTCAATCAACCCGAAGCCATTGCTTACATTAGT | 120 |
| Qy | 326 | GCCCATATTATGACGAGCGCGCTGGGAAAAAAACCGTTGCCACGCTTATGGAAGAG | 385 |

Db 1201 CATCACTCGGGCAAAATGGAACTTGTCACCCGCATGTTGCGCGCAGCTGAAGAGTATTCTAT 1260
QY 1466 GAATGTGGCTTTTGGGCAAAAGCAATAGCTCTAGCAAAAAAACAACCTGTAGAACAAAGT 1525
Db 1261 GAATGTGGCTTTTGGGCAAAAGCAATAGCTCTAGTAAAAAACAACCTGTAGAACAAAGT 1320
QY 1526 AGAAGCGGCGGCGATTGGTTTAAATTTGATGAAGACTGGGGCACAACACCAAGTGGAT 1585
Db 1321 AGAAGCGGCGGCGATTGGCTTTAAATTTGATGAAGACTGGGGCACAACACCAAGTGGAT 1380
QY 1586 CGATCACTGCTTGAGCGTGGCAGATCAATACGATGTGCAAGTTTGTATCCACACCGATAC 1645
Db 1381 CGATCACTGCTTGAGCGTGGCAGATCAATACGATGTGCAAGTTTGTATCCACACCGATAC 1440
QY 1646 AGTCAATGAGCGAGTTATGTAGATCACACCTTAAATGCAATGAACGCGGCGCCATCCA 1705
Db 1441 GGTCAATGAGCGAGTTATGTAGATCACACCTTAAATGCAATGAACGCGGCGCCATCCA 1500
QY 1706 TGCCTACCACTTGAAGGCGGGTGGAGGACACTCACCTGATGTTTATCAACCATGGCAGG 1765
Db 1501 TGCCTACCACTTGAAGGCGGGGAGGACACTCACCTGATGTTTATCAACCATGGCAGG 1560
QY 1766 CGAGCTCAATATTTACCTCTCTCCACACCCCACTATTCCCTATACCAATTAATACGGT 1825
Db 1561 CGAGCTCAATATTTACCTCTCTCCACACCCCACTATTCCCTATACCAATTAATACGGT 1620
QY 1826 TGCAGAACCTTAGACATGCTATGATGTCACACCTAGACAAACGATCCGCGAGGA 1885
Db 1621 TGCAGAACCTTAGACATGCTATGATGTCACACCTAGACAAACGATCCGCGAGGA 1680
QY 1886 TTTTACAAATTTCTCAAGCGGTATCGCGCCCGGCTCTATCGCGGCTGAAGATGTGCTCCA 1945
Db 1681 TCTCCAGTTTCCCAAGCGGTATCGCGCCCGGCTCTATCGCGGCTGAAGATGTGCTCCA 1740
QY 1946 TGATATGGGTGATCGCGATGAACAGCTCGGATTCGAAGCAATGGGCGTCAAGCGGA 2005
Db 1741 TGATATGGGTGATCGCGATGAACAGCTCGGATTCGAAGCAATGGGCGGCTGGGA 1800
QY 2006 AGTGATTCCTGAACTTGGCAGACTCGGATGAAGATTAAGAAATTTGGTAAGCTTCC 2065
Db 1801 AGTGATTCCTGAACTTGGCAGACTCGGATGAAGATTAAGAAATTTGGTAAGCTTCC 1860
QY 2066 TGAAGATGCAAGATGAATTTTCGCGATTAAGCGCTACATCTCCAAATACACTAT 2125
Db 1861 TGAAGATGTCAGATGAATGAACATTCGCGATCAACGCTATATCTCCAAATACACTAT 1920
QY 2126 CAACCCCGCTTTGACCCACCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGCAAGAT 2185
Db 1921 TAATCCCGCTTTGACCCATGGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGCAAGAT 1980
QY 2186 CGCGACTTGGTGTGGAATCTCGCTTTTGGCGGTAAACCCCAAAATCGTGATCAA 2245
Db 1981 CGCGCACTTGGTGTGGAATCTCGCTTTTGGTGTAAACCCCAAAATCGTGATCAA 2040
QY 2246 AGGCGGTATGTGCTTCTCTGAATGGCGGATTTAAGCGCTCTGTGCCACTCCCCA 2305
Db 2041 AGGCGGTATGTGCTTCTCTGAATGGCGGATTTAAGCGCTCTGTGCCACTCCCCA 2100
QY 2306 ACCGGTTTATACCGGAAATGTTTGGGATCACGGCAAGCGGAAATTTTACACACGAGCAT 2365
Db 2101 GCGGTTTATACCGGAAATGTTTGGGATCACGGCAAGCGGAAATTTTACACACGAGCAT 2160
QY 2366 CACTTTTGTTCAAAGTCGCTTATGAAATGGCGGTGAAGAAAGCTGGGCTTAGAGCG 2425
Db 2161 CACTTTTGTTCAAAGTCGCTTATGAAATGGCGGTGAAGAAAGCTTAGAGCG 2220
QY 2426 CCAAGTTTACCGGTCAAAAATCGCGGTAAACATCACCAAGAAAGACTTCAAGTTCAACGA 2485
Db 2221 CAAAGTGTACCGGTGAAGAAATCGCGCAACATCACTAAGAAAGACTTCAAGTTCAACAA 2280
QY 2486 CAAAACGGCAAAAATCACCGTGCATCCGAAACCTTCGAGGTCCTTGTAGATGCGAACT 2545
Db 2281 CAAAGCGGCGATATCACTGTGATCTTAACCTTCGAGGTCCTTGTAGATGCGAACT 2340

QY 2546 CTGCACCTCTAAACCCACCTCGCAAGTCCTCTAGCCAGCGCTACACTTTCTTCTAGGC 2605
Db 2341 CTGCACCTCTAAACCCCGCTCTGAAGTCCTCTAGCCCAACGCTACACTTTCTTCTAGGC 2400
QY 2606 ACAAT 2610
Db 2401 ACAAT 2405
RESULT 6
LOCUS AX356686 2405 bp DNA linear PAT 13-FEB-2002
DEFINITION Sequence 4 from Patent EP1176192.
ACCESSION AX356686
VERSION AX356686.1 GI:18674023
KEYWORDS
SOURCE Helicobacter felis
ORGANISM Helicobacter felis
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
Helicobacteraceae; Helicobacter.
REFERENCE
AUTHORS Kusters, J.G. and Cattoli, G.
TITLE Helicobacter felis vaccine
JOURNAL Patent: EP 1176192-A 4 30-JAN-2002;
Akzo Nobel N.V. (NL)
FEATURES
source Location/Qualifiers
1..2405
/organism="Helicobacter felis"
/mol_type="unassigned DNA"
/db_xref="taxon:214"
CDS
1..681
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD23257.1"
/db_xref="GI:18674024"
translation="MKLTPEQKFLVYAGEVARKKAEGLKLNQPEAIAYISAHIM
DEARRGKTVAELEMECHFLKDEVMGCVGNMVPDLGVEATFPDGTKLVTVMWPEP
DEHFKAGEVAFGCDKDIENAGKEVTELEVTNEGPKSLHVGSHFFETNKKALKKDRE
KAYGKLDIPSGNTLRIGAGQTRKVLQIPLGSGSKVIGMNGLVNNIADIERHKKALDK
AKSHGFIK"
692..2398
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD23258.1"
/db_xref="GI:18674025"
translation="MKMKQEVNVTYGTGDKVRLGDTDLWAEVHDYTTYGEELKF
GAKTIREGMGQNSPDENTLDLVTNAMIIDYTIKADIGIKNGKIHGICKAGNKD
MQGVSPHVVVGTEALAGEMLITAGGIDSHFLSPQFPPTALANGVTVMFGGCT
GPVDTNATITPGKWNLRMAAEEYSMMVFLGKGNSSSKQLVQVEAGAIGFK
LHEDGTPSAIDHCLSADEYDVQVCITHTDVNEAGYVDDTLNMGRAITHAYHIEG
AGGSHDPVITMAGELNLPSTPTPIPTINTVAEHLDMLTCHHLDKR IREDLQFS
QSRIRPGSIAAEDVLHDIGVMTATSSDSQAMGRAGEVTPRTWOTADKNKEFGKLPED
GADNDFRIKRYISKYITNPALTHGVSEYIGSVBEKIDLVVWNPAPFGVKKPIVIK
GMVWFSEMGDSNASVTTPQVITYREMFHGKAKFPDTSITFVSKVAENGVKELGL
ERKVLPIVKNCRNITKKDKFNKNKTAHITVDPKTFEVDGKLCITSKSPASEVPLAQRIT
FF"

ORIGIN

Query Match 74.9%; Score 2158.6; DB 6; Length 2405;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 2251; Conservative 0; Mismatches 154; Indels 0; Gaps 0;
QY 206 GTGAAACTCACACCAAGAGCAAGAAAGTCTTGTATATATGCGGCGCAAGTGGCT 265
Db 1 GTGAACTCACACCAAGAGCAAGAAAGTCTTGTATATATGCGGCGCAAGTGGCT 60
QY 266 AGAAGCGCAAGCAGAGGCTTAAAGCTCAACCAACCCGAAGCAATTCGCTTACATTAGT 325
Db 61 AGAAGCGCAAGCAGAGGCTTAAAGCTCAATCAACCCGAAGCCATTGCCTACATTAGT 120

| | | | |
|----|------|---|------|
| QY | 326 | GCCCATATATATGAGCGAAGCGCGCGTGGAAAAAACCCTTGCCAGCTTATGGAAGAG | 385 |
| Db | 121 | GGCCATATATATGAGCGAGCGCGCGTGGCAAAAAACCCTTGCTGAACTTATGGAAGAA | 180 |
| QY | 386 | TGCATGCACTTTTGNAAAAAGATGAGTAATGCCCCGGGGTGGTAAATATGGTTCCCGAT | 445 |
| Db | 181 | TGTATGCACTTTTGNAAAAAGATGAGGTGATGCCCGGGTGGGGAATATGGTCCCTGAT | 240 |
| QY | 446 | CTAGGTGTAGAAAGCAACCTTCTCTGATGTTACGAAAACCTTGTAACCTGTGAATTTGGCCCATC | 505 |
| Db | 241 | TTGGCGGTAGAGCCACTTTTCCCGGATGGCAACCAACTCGTAACCGTGAATTTGGCCCAT | 300 |
| QY | 506 | GAAACAGATGAGCACTTCAAGCGGGCGAAGTGAATTTGGTTGCGATAAAGACATCGAG | 565 |
| Db | 301 | GAACTGTATGAACACTTTAAAGCGGTGAAGTGAATTTGGCTGTGATAAAGACATTTGAG | 360 |
| QY | 566 | CTCAATGCAAGGCAAGAAGTAACCGAACTTTGAGGTTACTTAATGAAGGGCCTAAATCCTTG | 625 |
| Db | 361 | CTCAACGCGGGTAAAGGAAGTTACCGAGCTTGAAGTTTACCAACGAAGGACCTAAATCCTTG | 420 |
| QY | 626 | CATGTGGGTAGCCATTTCCACTTCTTTGAAAGCTTAAACAAGSCACTAAAAATTCGATCGTGAA | 685 |
| Db | 421 | CATGTGGGTAGCCATTTCCACTTCTTTGAAACCAACAGGCAATGAAATTCGATCGGAA | 480 |
| QY | 686 | AAAGCCTATGGCAACCGCCTAGATATTCCTCTGGCAACACGCTACGCAATTTGGGGCAGGA | 745 |
| Db | 481 | AAAGCCTATGGCAACCGCCTAGATATTCCTCTGGCAACACGCTACGCAATTTGGGGCAGGA | 540 |
| QY | 746 | CAAAACCGCAAGTGAGTTGATTCCTCTTTGGTGGCAGTAAAGAGTATGGCGATGAAC | 805 |
| Db | 541 | CAAAACCGTAAAGTGCAAGTTAATCCCTCTTTGGCGGTAGTAAAGAAAGTGAATTTGCGATGAAC | 600 |
| QY | 806 | GGCTTTGTCAATAACATCGCGATGAACGCGCATATAAACAATAAAGCGCTTGACAAAGGCGAAA | 865 |
| Db | 601 | GGCTTTGTGAATAAATTTGGGACGAAACGCGCATATAAACAATAAAGCGCTTGACAAAGGCGAAA | 660 |
| QY | 866 | TCTCACGGATTTATCAAGTAAGAGAGACTCCCATGAAAAATGAAAAACAAGAAATATGTAAA | 925 |
| Db | 661 | TCTCACGGATTCATCAAGTAAGAGAGACTCCCATGAAAAATGAAAAACAAGAGTATGTAAA | 720 |
| QY | 926 | TACCTACGGAACCCACCAAGGCGATAAAGTGCCTTAGGAGATACCGATCTTTGGGCGAGA | 985 |
| Db | 721 | CACCTACGGAACCCACCAAGGCGATAAAGTGCCTTAGGAGATACCGATCTTTGGGCGAGA | 780 |
| QY | 986 | AGTAGAACATGACTATACCACTCTCGCGAAGAACTTAAATTTGGCGCGGTAAACTAT | 1045 |
| Db | 781 | AGTAGAACATGACTATACCACTTTATGGCGAAGAGCTCAAAATTTGGCGCGGTAAACTAT | 840 |
| QY | 1046 | CGGTGAGGGTATGGGTACAGAGCAATAGCCCTGATGAAAAACACCTAGATTTAGTCATCAC | 1105 |
| Db | 841 | CGGTGAGGGTATGGGTACAGAGCAATAGCCCGATGAAAAACACCTTAGATTTAGTCATCAC | 900 |
| QY | 1106 | TAAACGCGATGATPATCGACTACACCGGATTTTCAAAAGCCGACATTTGGGATTTAAAAACGG | 1165 |
| Db | 901 | CAACGCGATGATPATCGACTACACCGGATTTTAAAGCCGACATTTGGTATTTAAAAATGG | 960 |
| QY | 1166 | CAAAATCCATGGCATTTGGCAAGCAGGAACAAGGACATCAAGATGGCGTAAAGCCCTCA | 1225 |
| Db | 961 | CAAAATCCATGGTATTTGGCAAGCAGGGAACAAGACATCAAGATGGCGTAAAGCCCTCA | 1020 |
| QY | 1226 | TATGGTCTGGGTGTGGGCAACAAGCACTAGCAGAGGGAAGGTATGATTTATTCACGCTGG | 1285 |
| Db | 1021 | TATGGTCTGGGTGTGGGCAACAAGCACTAGCAGGGAAGGTATGATTTATTCACGCTGG | 1080 |
| QY | 1286 | GGGAATCGATTTCAACACCCACTTCTCTTTCTCCACAAACAATTCCTACCGCTCTAGCCAA | 1345 |
| Db | 1081 | GGGGATCGATTTGCGACACCCACTTCTCTCTCCCAACAATTCCTACCGCTCTAGCCAA | 1140 |
| QY | 1346 | TGGCGTTACAAACATTTTGGAGGCGGCAAGGTCCTGTAGATGGCAAGAAATGCACTAC | 1405 |
| Db | 1141 | TGGTGTTAACAACCATGTTTGGAGGTGGCAAGGTCCTGGTATGTCGCAAGAAATGCACTAC | 1200 |
| QY | 1406 | TATCACTCCGGGCAAAATGGAACTTGACCGCATGTTGGCGCGACAGAGATTTCTAT | 1465 |
| Db | 1201 | CATCACTCCGGGCAAAATGGAACTTGACCGCATGTTGGCGCGACGTGAAGAGTATTTCTAT | 1260 |
| QY | 1466 | GAATGTGGGCTTTTGGGCAAAAGCAATAGCTCTAGCAAAAAACAACCTTTGTAGAACAACT | 1525 |
| Db | 1261 | GAATGTGGGCTTTTGGGCAAAAGCAATAGCTCOAGTAAAAACAACCTCTGTAGAACAACT | 1320 |
| QY | 1526 | AGAAGCGGGCGCGATTTGGTTTAAATTTGCATGAAGACTGGGGGCACAAACCAAGTGCAT | 1585 |
| Db | 1321 | AGAAGCGGGCGCGATTTGGCTTTAAATTTGCATGAAGACTGGGGGCACAAACCAAGTGCAT | 1380 |
| QY | 1586 | CGATCACTGCTTGAGCGGTGGCAGATGAATAGGATGTGCAAGTTTGTATTCACACCGCATAC | 1645 |
| Db | 1381 | CGATCACTGCTTGAGCGGTAGCAGATGAATAGGATGTGCAAGTTTGTATTCACACCGCATAC | 1440 |
| QY | 1646 | AGTCAATGAGGCAAGTTATGTAGATGACACCTTAAATGCAATGAACGGCGCGCCATCCA | 1705 |
| Db | 1441 | GGTCAATGAGGCAAGTTATGTAGATGACACCTTAAATGCGATGAACGGCGCGCCATCCA | 1500 |
| QY | 1706 | TGCTTACCACTTGAGGGAGCGGTGGAGGACACTCACCTGATGTTATCAACCATGGCAGG | 1765 |
| Db | 1501 | TGCTTACCACTTGAGGGAGCGGTGGAGGACACTCACCTGATGTTATCAACCATGGCAGG | 1560 |
| QY | 1766 | CGAGCTCAATATTTCTACCTCTCCACCAACCCCACTATTTCCCTATATCAATTAATACGGT | 1825 |
| Db | 1561 | CGAGCTCAATATTTCTACCTCTCCACCAACCCCACTATTTCCCTATATCAATTAATACGGT | 1620 |
| QY | 1826 | TGCAGAACACTTAGACATGCTCATGACATGCCACCACTAGACAAACGCATCCCGGAGGA | 1885 |
| Db | 1621 | TGCAGAACACTTAGACATGCTCATGACATGCCACCACTAGACAAACGCATCCCGGAGGA | 1680 |
| QY | 1886 | TTTACAAATTTTCTCAAAAGCGGTATCCGCGCGGCTCTATCCGCGCTGAAGATGTGCTCCA | 1945 |
| Db | 1681 | TTTCCAGTTTTCCTCAAAAGCGGTATCCGCGCGGCTCTATTCGCGCTGAAGATGTGCTCCA | 1740 |
| QY | 1946 | TGATATGGGTGTGATCCGATGAACAAGCTCGGATTCGCAAGCAATTTGGGCGGTGCAAGGA | 2005 |
| Db | 1741 | TGATATGGGTGTGATCCGATGAACAAGCTCGGATTCGCAAGCAATTTGGGCGGTGCAAGGA | 1800 |
| QY | 2006 | AGTGATTTCTCGAACTTTGGCAGACTGCGGATGAAGATAAGAAATTTGTAAGCTTCC | 2065 |
| Db | 1801 | AGTGATTTCTAGAACTTTGGCAAACTGCAGACAAGAAATTTAAAGAAATTTGTAAGCTTCC | 1860 |
| QY | 2066 | TGAAGATGGCAAAAGATAAACGATTAATTTCCGATTAAGCGCTACATCTCCAAATACACTAT | 2125 |
| Db | 1861 | TGAAGATGGTGCAGATTAATGCAAACTTTCCGATTCAAAGCTATATCTCCAAATACACTAT | 1920 |
| QY | 2126 | CAACCCCGCTTTTGACCCACCGCGTGAAGTATATCGGCTCTGTGGAAGAGGCGCAAGAT | 2185 |
| Db | 1921 | TAATCCCGCTTTTGACCCATGGCGTGAGGAGTATATCGGCTCTGTGGAAGAGGCGCAAGAT | 1980 |
| QY | 2186 | CGCGCATTTGGTGTGGAAATCTCGCTTTTGGCGGTAAAAACCAAAATCGTATCAA | 2245 |
| Db | 1981 | CGCGCATTTGGTGTGGAAATCTCGCTTTTGGTGTAAAAACCAAAATCGTATCAA | 2040 |
| QY | 2246 | AGGCGGTATGGTGTCTCTGAAATGGGCGATTTCAACGCTCTGTGCGCCACTCCCA | 2305 |
| Db | 2041 | AGGCGGTATGGTGTCTCTGAAATGGGCGATTTCAACGCGCTGTGCGCCACACTCA | 2100 |
| QY | 2306 | ACCGGTTTATTTACCGGAAATTTTGGGCACTCAAGGCAAGGCGAAATTTGACACCAAGAT | 2365 |
| Db | 2101 | GCGGTTTATTTACCGGAAATTTTGGGCACTCAAGGCAAGGCGAAATTTGACACCAAGAT | 2160 |
| QY | 2366 | CACTTTGTTCCTCAAGTTCGCTATGAAAAATGGCGGTGAAAGAAAGCTGGGCTTAGAGCG | 2425 |
| Db | 2161 | CACTTTGTTCCTCAAGTTCGCTATGAAAAATGGCGGTGAAAGAAAGCTAGGCTTAGAGCG | 2220 |
| QY | 2426 | CCAAGTTCTACCGGTCAAAACTTCGCGTAAACATCACCAAGAAAGACTTCAAGTTCAACGA | 2485 |
| Db | 2221 | CAAGGTCTACCGGTCAAAACTTCGCGCAACATCACTAAGAAAGACTTCAAAATCAACAA | 2280 |
| QY | 2486 | CAAAACGGCAAAATCAACCGTCAATCCGAAACCTTCGAGGTCTTTGTAGATGGCAAACT | 2545 |

| | | | | |
|---|---|--|-----|------------------------|
| Db | 2281 | CAGACGGCGCATATCACTGTCGATCTCTAAACCTTCGAGGTCTTTGTAGATGGCAAACT | | 2340 |
| Qy | 2546 | CTGCACCTCTAAACCCACCTCGCAAGTGCCTCTAGCCACGCTACACTTTCTTCTAGGC | | 2605 |
| Db | 2341 | CTGCACCTCTAAACCCGCTCTGAAGTGCCTCTAGCCCAAGCTACACTTTCTTCTAGGC | | 2400 |
| Qy | 2606 | ACAAAT 2610 | | |
| Db | 2401 | ACAAAT 2405 | | |
| RESULT 7 | | | | |
| LOCUS | BD185305 | 2407 bp | DNA | linear PAT 17-JUN-2003 |
| DEFINITION | Helicobacter felis vaccine. | | | |
| ACCESSION | BD185305 | | | |
| VERSION | BD185305.1 | GI:31877505 | | |
| KEYWORDS | JP 2002355054-A/4. | | | |
| SOURCE | Helicobacter felis | | | |
| ORGANISM | Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Helicobacter. | | | |
| REFERENCE | 1 (bases 1 to 2407) | | | |
| AUTHORS | Kusters,J.G. and Cattoli,G. | | | |
| TITLE | Helicobacter felis vaccine | | | |
| JOURNAL | Patent: JP 2002355054-A 4 10-DEC-2002; | | | |
| COMMENT | AKZO NOBEL NV | | | |
| OS | Helicobacter felis | | | |
| | PN | JP 2002355054-A/4 | | |
| | PD | 10-DEC-2002 | | |
| | PF | 16-JUL-2001 JP 2001214711 | | |
| | PR | 17-JUL-2000 EP 00202565.8 | | |
| PI | JOHANNES GERARDUS KUSTERS,GIOVANNI CATTOLI | | | |
| PC | C12N15/09,A61K38/00,A61K39/106,A61K39/118,A61K39/12,A61K39/175, PC | | | |
| PC | A61K39/23, | | | |
| | PC | A61K39/235,A61K39/39,A61K39/395,A61P1/04,A61P31/04,C12N1/15, | | |
| | PC | C12N1/19, | | |
| PC | C12N1/21,C12N5/10,C12N9/80,C12Q1/68,G01N33/15,G01N33/50,G01N33/ PC | | | |
| PC | 53, | | | |
| | G01N33/53,G01N33/566,G01N33/569// | (C12N9/80,C12R1:01),(C12Q1/68, PC | | |
| | C12R1:01), | | | |
| | PC | C12N15/00,C12N5/00,A61K37/02 | | |
| | CC | Helicobacter felis vaccine | | |
| Key | Location/Qualifiers | | | |
| FT | (2)..(682) | | | |
| FT | (693)..(2399). | | | |
| FEATURES | Location/Qualifiers | | | |
| source | 1..2407 | | | |
| ORIGIN | /organism="Helicobacter felis" | | | |
| | /mol_type="genomic DNA" | | | |
| | /db_xref="taxon:214" | | | |
| Query Match 74.9%; Score 2158.6; DB 6; Length 2407; | | | | |
| Best Local Similarity 93.6%; Pred. No. 0; | | | | |
| Matches 2251; Conservative 0; Mismatches 155; Indels 0; Gaps 0; | | | | |
| Qy | 206 | GTGAACCTACACCCAAAGCAGGAGTCTTAAAGCTCAACCAACCCGAGCCATTGCTTACATTAGT | 325 | |
| Db | 2 | GTGAACCTACACCCAAAGCAGGAGTCTTAAAGCTCAACCAACCCGAGCCATTGCTTACATTAGT | 61 | |
| Qy | 266 | AGAAAGCGCAAGCAGGAGTCTTAAAGCTCAACCAACCCGAGCCATTGCTTACATTAGT | 325 | |
| Db | 62 | AGAAAGCGCAAGCAGGAGTCTTAAAGCTCAACCAACCCGAGCCATTGCTTACATTAGT | 121 | |
| Qy | 326 | GCCCATATTATGGACGAAGCGCCGCTGGAAAAAAACCGTTGCCAGCTTATGGAAGAG | 385 | |
| Db | 122 | GCCCATATTATGGACGAAGCGCCGCTGGAAAAAAACCGTTGCCAGCTTATGGAAGAG | 181 | |
| Qy | 386 | TGCATGCACCTTTTGGAAAAAGATGAAGTAAATGCCCGGGGTGGGTAATATGTTGGTCCCGAT | 445 | |

| | | | |
|----|------|---|------|
| Db | 182 | TGATGACACTTTTGGAAAAAGACGAGGTATGCCCGGTGTGGGAAATATGTCCTGAT | 241 |
| Qy | 446 | CTAGTGTAGAAAGCCACCTTTCTCTGATGGTACGAAACTTGTAACTGTGAAATTTGGCCCATC | 505 |
| Db | 242 | TTAGGCGTGGAGACTCTTTTCCGATGGCACCAACTCGTAACCGTGAATTTGGCCCATC | 301 |
| Qy | 506 | GAACAGATGAGCACTTCAAAAGCGGGCGAAGTGAATTTGGTTGGATAAAGACATCGAG | 565 |
| Db | 302 | GAACCGGATGAACACTTCAAAAGCGGGCGAAGTCAAAATTTGGCTGTGATAAAGACATTGAA | 361 |
| Qy | 566 | CTCAATGCGCAAAAGTAACCGAATCTGAGGTACTAATGAAGGGCCTAAATCCTTGG | 625 |
| Db | 362 | CTCAACGAGGTAAGGAAGTTACCGAATCTGAGGTACTAATGAAGGGCCTAAATCCTTGG | 421 |
| Qy | 626 | CATGTGGGTAGCCATTTCCACTTCTTGAAGCTTAACAGGCACTAAAAATTCGATCGTGA | 685 |
| Db | 422 | CATGTGGGTAGCCATTTCCACTTCTTGAAGCTTAACAGGCACTAAAAATTCGATCGTGA | 481 |
| Qy | 686 | AAAGCCTATGCAAAACGCTAGATATTCCTCTCGCAACACGCTACGATTTGGGGCAGA | 745 |
| Db | 482 | AAAGCCTATGCAAAACGCTAGATATTCCTCTCGCAACACGCTACGATTTGGGGCAGA | 541 |
| Qy | 746 | CAAAACCGCAAGTGCAGTTGATTCCTCTTGGTGGCAGTAAAAAGTGAATTTGGCATGAA | 805 |
| Db | 542 | CAAAACCGTAAAGTGCAGTTAATCCCTCTTGGCGGCACTAAAAAGTGAATTTGGCATGAA | 601 |
| Qy | 806 | GGGCTTGTGAATAACATCGCGGATGAAGCCATTAACATAAAGCGCTTGACAAGCGAAA | 865 |
| Db | 602 | GGGCTTGTGAATAATAATGCAAGTGAAGTGAAGCCATTAACATAAAGCGCTTGACAAGCGAAA | 661 |
| Qy | 866 | TCTCAGCGATTATCAAGTGAAGGAGACTCCCATGAAATGAAAAACAAAGATATGTAAA | 925 |
| Db | 662 | TCTCAGCGATTATCAAGTGAAGGAGACTCCCATGAAATGAAAAACAAAGATATGTAAA | 721 |
| Qy | 926 | TACCTACGGACCCCAAGGCGGATTAAGTGCCTTAGGAGATACCGATTTTGGGCGAG | 985 |
| Db | 722 | TACCTACGGACCTACCAAGGCGGATTAAGTGCCTTAGGAGATACCGATTTTGGGCGAG | 781 |
| Qy | 986 | AGTAGAATGATGATATACCATGATGCGAAGAACTTAAATTTGGCGGGGTAAAACTAT | 1045 |
| Db | 782 | AGTAGAATGATGATATACCATGATGCGAAGAACTTAAATTTGGCGGGGTAAAACTAT | 841 |
| Qy | 1046 | CCGTGAGGTTATGGTCAGAGCAATAGCCCTGATGAACCAACCCCTAGATTTAGTCATCAC | 1105 |
| Db | 842 | CCGTGAGGTTATGGTCAGAGCAATAGCCCTGATGAACCAACCCCTAGATTTAGTCATCAC | 901 |
| Qy | 1106 | TAAACCGGATGATTAACGACTACACCGGATTTACAAAGCCGACATTTGGGATTTAAAAACGG | 1165 |
| Db | 902 | CAACCGGATGATTAACGACTACACCGGATTTACAAAGCCGACATTTGGGATTTAAAAACGG | 961 |
| Qy | 1166 | CAAAATCCATGGCATTTGGCAAGCGGAGGAAACAAGGACATGCAAGATGGCGTGAAGCCCTCA | 1225 |
| Db | 962 | CAAAATCCATGGCATTTGGCAAGCGGAGGAAACAAGGACATGCAAGATGGCGTGAAGCCCTCA | 1021 |
| Qy | 1226 | TATGGTCTGGGTGTGGGCACAGAGCACTAGCAGGGAAGGTATGATTTATCCGCTGG | 1285 |
| Db | 1022 | TATGGTCTGGGTGTGGGCACAGAGCACTAGCAGGGAAGGTATGATTTATCCGCTGG | 1081 |
| Qy | 1286 | GGGAATCGATTACACACCCCACTTCTTCTTCCCAACAAATTCCTACCGCTCTAGCCCAA | 1345 |
| Db | 1082 | GGGAATCGATTACACACCCCACTTCTTCTTCCCAACAAATTCCTACCGCTCTAGCCCAA | 1141 |
| Qy | 1346 | TGGCGTTTCAACCACTGTTTGGAGCGGCACAGGTCCTGTAGATGGCAAGATTCGACTAC | 1405 |
| Db | 1142 | TGGCGTTTCAACCACTGTTTGGAGCGGCACAGGTCCTGTAGATGGCAAGATTCGACTAC | 1201 |
| Qy | 1406 | TATCAGCTCCGGGCAAAATGGAACCTTGACCGCATGTTGGCGGCGAGCAGAGATTTCTAT | 1465 |
| Db | 1202 | CATCAGCTCCGGGCAAAATGGAACCTTGACCGCATGTTGGCGGCGAGCAGAGATTTCTAT | 1261 |
| Qy | 1466 | GAATGTGGGCTTTTGGCAAGGCAATAGCTCTAGCAAAAAAACAATTTGTAGAAACAAGT | 1525 |

| | | | | |
|----|------|---|--------------------------------------|------|
| Db | 1262 | GAATGTGGGCTTTT | TGGGCAAAAGCGAATAGCTCCAGTAAAAAACAAC | 1321 |
| Qy | 1526 | AGAAAGCGGGCGCGAT | TGGTTTTAAATTTGCATGAAGACTGGGGCAACAACA | 1585 |
| Db | 1322 | AGAAAGCGGGCGCGAT | CGGCTTTAAATTTGCATGAAGACTGGGGCAACAAC | 1381 |
| Qy | 1586 | CGATCACTGCTTGAGCGGTGCGAGATGAATACGATGTGCAAGTTGTATCCACACCGATAC | 1645 | |
| Db | 1382 | CGATCACTGCTTGAGCGGTAGCAGATGAATACGATGTGCAAGTTGTATCCACACCGATAC | 1441 | |
| Qy | 1646 | AGTCAATGAGGCAAGTTATCTAGATGACACCCCTAAATGCAATGAACGGGCGCGCATCCA | 1705 | |
| Db | 1442 | GGTCAATGAGGCAAGTTATCTAGATGACACCCCTAAATGCAATGAACGGGCGCGCATCCA | 1501 | |
| Qy | 1706 | TGCTACACCAATTTGAGGGAGCGGTGAGGACACTCACTGATGTTATCACCATGGCAGG | 1765 | |
| Db | 1502 | TGCTACCAATTTGAGGGAGCGGTGAGGACACTCACTGATGTTATCACCATGGCAGG | 1561 | |
| Qy | 1766 | CGAGCTCAATATTCTACCTTCTCCACCAACCCCGACATTTCCCTATACCAATTAATACGTT | 1825 | |
| Db | 1562 | CGAGCTCAATATTCTACCTTCTCCACCAACCCCGACATTTCCCTATACCAATTAATACGTT | 1621 | |
| Qy | 1826 | TGCAGAACCTTAGACATGCTCATGACATGCCACCACTAGACAAACGCAATCCGCGAGGA | 1885 | |
| Db | 1622 | TGCAGAACCTTAGACATGCTCATGACCTGCCACCACTAGATAAAGCGCAATCCGCGAGGA | 1681 | |
| Qy | 1886 | TTTACAATTTTCTCAAAAGCGGTATCCGCCCGGGCTCTATCGGGCTCAAGATGTGCTCCA | 1945 | |
| Db | 1682 | TTTACAATTTTCCAAAGCGGTATCCGCCCGGGCTCTATCGCGCTGAAAGATGTGCTCCA | 1741 | |
| Qy | 1946 | TGATATCGGTGTGATCCGATGACAAGCTCGGATTCGCAAGCAATGGGGCGTGCAGCGA | 2005 | |
| Db | 1742 | TGATATTTGGGTGTATCGCATGACAAGCTCGGATTCGCAAGCAATGGGGCGTGCAGCGA | 1801 | |
| Qy | 2006 | AGTGATTTCTCGAATTTGGCAGACTCGGATAGAAATAAAAGAAATTTGGTAAAGCTTCC | 2065 | |
| Db | 1802 | AGTGATTTCTCGAATTTGGCAGACTCGGATAGAAATAAAAGAAATTTGGTAAAGCTTCC | 1861 | |
| Qy | 2066 | TGAAGATGGCAAGATAACGATAATTTCCGATTAAAGCGCTACATCTCCAAATACACTAT | 2125 | |
| Db | 1862 | TGAAGATAGTCAGATAACGACAATTTCCGTTATCAAGCGTACATCTCCAAATACACTAT | 1921 | |
| Qy | 2126 | CAACCCCGCTTTGACCCACGCGTGAAGCTATATCGGCTCTGTGGAAGAGGGCAAGAT | 2185 | |
| Db | 1922 | TAACCCCGCTCTAACCCCATGGGTAAAGCAGTATATCGGCTCTGTGGAAGAGGGCAAAAT | 1981 | |
| Qy | 2186 | CGCCGACTTGGTGTGTGGAATCTCGCTTTTTTTGGCGTAAACCCCAAAATCTGTGATCAA | 2245 | |
| Db | 1982 | CGCTGATTTGGTGTGTGGAATCTCGCTTTTTTTTGGTGTGAAACCTAAGATTTGTGATCAA | 2041 | |
| Qy | 2246 | AGGCGGTATGGTGTCTTCTGAAATGGGCGATTCTAAACGCTCTGTGCCCTACCTCCCA | 2305 | |
| Db | 2042 | AGGCGGTATGGTGTCTTCTGAAATGGGCGATCTCAAGCGTCTGGTCCCTACCCCTCA | 2101 | |
| Qy | 2306 | ACCGGTTTATTACCGCGAAATGTTTGGGCATCACGGCAAGCGCAAAATTTGACACCAAGAT | 2365 | |
| Db | 2102 | GCCGGTTATTACCGCGAAATGTTTGGGCATCACGGCAAGCGCAAAATTTGACACCAAGAT | 2161 | |
| Qy | 2366 | CACTTTGTTTTCCAAAGTCGCTATGAAATGGCGTGAAGAAAGCTGGGCTTAGAGCG | 2425 | |
| Db | 2162 | CACTTTGTTTTCCAAAGTCGCTATGAAATGGGCGTGAAGAAAGAACTTAGAGCG | 2221 | |
| Qy | 2426 | CCAAAGTCTACCGGTCAAAACTCCGCTAACATCACCAAGAAAGACTTCAAGTTCAACGA | 2485 | |
| Db | 2222 | CAAGGTGCTACCGGTGAAAAACTCCCGCAACATCACTAAGAAAGACTTCAAAATTCACAA | 2281 | |
| Qy | 2486 | CAAAACGGCAAAATTCACCGTCGATCCGAAACCTTCAGAGTCTTTGTAGATGGCAAACT | 2545 | |
| Db | 2282 | CAAGCGGCGCATATCACTGTGATCCTTAAACCTTCGAGGTCTTTGTAGATGGCAAACT | 2341 | |
| Qy | 2546 | CTGACCTCTTAAACCCACTCGCAAGTGCCTCTTAGCCAGCGCTACACTTTCTTTTAGGC | 2605 | |
| Db | 2342 | CTGACCTCTTAAACCCGCTCTGAAGTGCCTCTAGCCAGCGCTACACTTTCTTTCTAGGC | 2401 | |

```

Qy 2606 ACAATG 2611
    |||||
Db 2402 NCAATG 2407

RESULT 8
LOCUS AX356692 2407 bp DNA linear PAT 13-FEB-2002
DEFINITION Sequence 10 from Patent EP1176192.
ACCESSION AX356692
VERSION AX356692.1 GI:18674029
KEYWORDS
SOURCE
  ORGANISM
    Helicobacter felis
    Helicobacter felis
    Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
    Helicobacteraceae; Helicobacter.
REFERENCE
  1 Kusters, J.G. and Cattoli, G.
    Helicobacter felis vaccine
    Patent: Bp 1176192-A 10 30-JAN-2002;
    Akzo Nobel N.V. (NL)
FEATURES
  source
    1..2407
      /organism="Helicobacter felis"
      /mol_type="unassigned DNA"
      /db_xref="taxon:214"
    2..-682
      /note="unnamed protein product"
      /codon_start=1
      /transl_table=11
      /protein_id="CAD23262.1"
      /db_xref="GI:18674030"
      /translations="MKLTPEQSKFLLYVAGEVARKRKAEGKLINOPEAITYAISAHM
      DEARRKKTVAELMESCMHFLKDEYMPGVGMVNDLGEATFPGDTKLVTVNMPIEP
      DEFHKAQVFGCDKDIELNAGREVTELEVTEGPKSLHVGSHFHFPEANKALKFRDE
      KAYGKRLDIPSGNLTIRIGAGQTRKVLQIPLUGGSKKVIWMGNLVNNIADBRHRHKALEK
      AKSHGFIK"
    693..2399
      /note="unnamed protein product"
      /codon_start=1
      /transl_table=11
      /protein_id="CAD23262.1"
      /db_xref="GI:18674031"
      /translations="MMKKQEVYNTYGPITTKDKVRLGDTDLMAEVSHDYTTYGEELKF
      GAKTIREGQSNSPDENTLDLIVTNMIIDYTGIVKADIGIKNGKIHGIGKAGNKO
      MQGDVSPHVVVGVTGEBALAGEHMIITAGDISHTHFLSQOQFTALANGVTTWFGGT
      GPDVGTTPDAIDCHLSVADYDQVCIHTDTVNEAGYDDTLNANMGRAIHAYHTEG
      ACGHGPDSIAIDAGELNIIPLSSSTPTPIPTINTVAEHLDMLTCHHLDKRIREDLQF
      QSRIRPGSIAAEVDLHDIQVIAMTSDSOAMGRAGEVIPTWQIADPNKKEFKCLPED
      SGNDRPFRMGIKSYITNPALTHGVSEYIGSVEEGKIADLVVNPAPFGVKPKIVIK
      GMMVFSMGDSNASPTPOQVYIREMFGHHGKAFDTSITFVSKVAYENGVEKKGJL
      ERYKLPVKNCRNRTTKKDFPNKNKTAHIITVDPKTFEYFVDGKCLCTSKPASEVPLAQRYT
      FF"

ORIGIN
  Query Match 74.9%; Score 2158.6; DB 6; Length 2407;
  Best Local Similarity 93.6%; Pred. No. 0;
  Matches 2251; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

Qy 206 GTGAAACTCACACCAAGACGAGAAAGTCTTGTGTATATTATGCGGCGCAAGTGGCT 265
    |||||
Db 2 GTGAACTCACCCAAAGACGAGAAAGTCTTGTGTATATTATGCGGCGCAAGTGGCT 61
    |||||

Qy 266 AGAAAGCGCGAAAGCAGAGCGGCTTAAAGCTCAACCAACCCGCAAGCCATTGCTTACATTAGT 325
    |||||
Db 62 AGAAAGCGCGAAAGCGGAGGCTTAAAGCTCAACCAACCCGCAAGCCATTGCTTACATTAGT 121
    |||||

Qy 326 GCCCATATTATGACGAGCGCGCGTGGGAAAAAAMCCGTTGCCAGCTTATGGAAGAG 385
    |||||
Db 122 GCCCATATTATGACGAGCGCGCGTGGGAAAAAAGCCGTTGCGGAACCTTATGGAAGAG 181
    |||||

```

QY 386 TGCATGCACTTTTGTAAAAAAGATGAAGTAATGCCCCGGGGTGGTAATATATGTTCCCGAT 445
Db 182 TGTATGCACTTTTGTAAAAAAGACAGGTGATGCCGGTGTGGGAAATATGGTCCCTCAT 241
QY 446 CTAGGTGTAGAAGCCACTTTCTGTAGTGTACGAACTTGTACTGTGAATTTGGCCCATC 505
Db 242 TTAGGGGTGAAGCTACTTTTCCGATGGCACCAACTCGTAACCGTGAATTTGGCCCATC 301
QY 506 GAACCCAGATGAGCACTTCAAAGCGGGCGAAAGTGAATTTGGTTGCGATAAAGACATCGAG 565
Db 302 GAACCCGATGACACTTCAAAGCGGGCGAAGTCAATTTGGCTGTGATAAAGACATTTGAA 361
QY 566 CTCAATGCAAGCAAGAAAGTAACCGAACTTTGAAGTTTACTAATGAAGGCCCTTAATCTCTTG 625
Db 362 CTCAAACGAGTGAAGGAAGTTACCGAACTAGAAGTTACCAACGAAGGACCTAAATCTCTTG 421
QY 626 CATGTGGGTAGCCATTTCACACTTCTTTGAAGCTTAAAGGCACTTAAATTCGATCGTGAA 685
Db 422 CATGTGGGTAGCCATTTCACACTTCTTTGAAGCCAAACGAAGGCACTTGAATTCGATCGGGAA 481
QY 686 AAAGCCTATGCGAAACCGCTTAGATATTCCCTCTGGCAACAGCTACGCAATTTGGGGCAGGA 745
Db 482 AAAGCCTATGCGAAACCGCTTAGATATTCCCTCTGGCAACAGCTACGCAATTTGGGGCAGGA 541
QY 746 CAAACCCGCAAGTGCAAGTTGATTCCTCTTTGGTGGCAGTAAAAAGTGATTTGGCATGAAC 805
Db 542 CAAACCCGTAAGTGCAAGTTAATCCCTCTTTGGCGCAGTAAAAAGTGATTTGGCATGAAC 601
QY 806 GGCCTTGTGAATTAACATCGCGGATGAACGCGCATAAACATTAAGCGCTTTGAACAGGCGAA 865
Db 602 GGCCTTGTGAATTAATTTGCAGATGAACGCGCATAAACATTAAGCGCTTTAGAAAAAGCAAA 661
QY 866 TCTCAGGATTTATCAAGTAAGAGACTCCCATGAAATGAAGAAACCAAGAAATGTGAAA 925
Db 662 TCTCAGGATTTATCAAAATAGGAGACTCCCATGAAATGAAGAAACCAAGAGTATGAAA 721
QY 926 TACCTACGGACCCACAAAGGCGATAAGTGGCTTAGGAGATACCGCATTTTGGGCAGA 985
Db 722 TACCTACGGACCTTACCACAGCGCAAAAGTGGCTTAGGAGATACCGCATTTTGGGCAGA 781
QY 986 AGTAGAACATGATATACCACTTATGGCGAAGAACTTAAATTTGGCGCGGTAAAACTAT 1045
Db 782 AGTAGAACATGATATACCACTTATGGCGAAGAGCTCAATTTTGGCGCGGTAAAACTAT 841
QY 1046 CGGTAGGGTATGGGTGAGAGCAATAGCCCTGTATGAAAAACCCCTAGATTAGTCATCAC 1105
Db 842 CGGTAGGGCATGGGTGAGAGCAATAGTCAGATGAAAAACCCCTAGATTAGTCATCAC 901
QY 1106 TAACGCGATGATTTATCGACTTACACCGGGATTTACAAAGCCGACATTCGGGATTTAAAAACGG 1165
Db 902 CAAACGCGATGATTTATGACTTACACCGGGATTTACAAAGCCGACATTCGGCATTTAAAAATGG 961
QY 1166 CAAAAATCCATGGCAITTTGGCAAGGCAGGAAACAAAGGACATGCAAGATGGCGTAAAGCCCTCA 1225
Db 962 CAAAAATCCATGGCAITTTGGCAAGGCAGGAAACAAAGGACATGCAAGATGGCGTAAAGCCCTCA 1021
QY 1226 TATGGTGTGGGTGAGGCAAGAGCACTAGCAGGGGAAGGTATGATTAATTCACGCTGG 1285
Db 1022 TATGGTGTGGGTGAGGCAAGAGCACTAGCAGGGGAAGGTATGATTAATTCACGCTGG 1081
QY 1286 GGGAAATCGATTACACACCCACTTCTTCTCCACAACTTCCCTACCGCTCTAGCCAA 1345
Db 1082 GGGAAATCGATTACACACCCACTTCTTCTCCACAACTTCCCTACCGCTCTAGCCAA 1141
QY 1346 TGGCGTTACAAACCATGTTTGGAGCGGCACAGGTCTCTGTAGATGGCACGAATGCGACTAC 1405
Db 1142 TGGCGTTACAAACCATGTTTGGCGGTGGCACAGGTCCGGTAGATGGCACGAATGCGACTAC 1201
QY 1406 TATCACTCGGGCAATTTGAACTTTGCAACCGCATTTGTCGCGGACAGAGAGTATTTCTAT 1465
Db 1202 CATCACTCGGGCAATTTGAACTTTGCAACCGCATTTGTCGCGGACAGTGAAGAGTATTTCTAT 1261
QY 1466 GAATGTGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACCACTTGTAGAACAAAT 1525

Db 1262 GAATGTGGGCTTTTGGGCAAGGCAATAGCTCCAGTAAAAAACAACTTGTAGAACAAAT 1321
QY 1526 AGAAGCGGGCCGATTTGGTTTAAATTCGATGAAGACTTGGGGCAAAACCAAGTGGCAT 1585
Db 1322 AGAAGCGGGCCGATCGGCTTTAAATTCGATGAAGACTTGGGGCAAACTTCAAGTGCAT 1381
QY 1586 CGATCACTGCTTGAAGCGTGGCAGATGAATAAGATGCAAGTTTGTATCCACACCGATAC 1645
Db 1382 CGATCACTGCTTGAAGCGTGGCAGATGAATAAGATGCAAGTTTGTATCCACACCGATAC 1441
QY 1646 AGTCAATGAGCAGGTTATGTAGATGACACCCCTAAATGCAATGAACGGCGGCCATCCA 1705
Db 1442 GGTCAATGAGCAGGTTATGTAGATGACACCCCTGAATGCGATGAACGGCGGCCATCCA 1501
QY 1706 TGCCTACCACTTGAAGGAGCGGGTGGAGGACACTCACTGATGTTTATCACCATTGGCAGG 1765
Db 1502 TGCCTACCACTTGAAGGAGCGGGCGAGGACACTCACTGATGTTTATCACCATTGGCAGG 1561
QY 1766 CGAGCTCAATATTCTTACCTCTCCACACCCCACTATTCCTTATACCATTAATACGGT 1825
Db 1562 CGAGCTCAATATTCTACCTCTCCACAAACCCCACTATCCCTATACCATTAATACGGT 1621
QY 1826 TGCAGAACACTTGAACATGCTCATGACATGCCACCACTAGACAAACGGATCCGCGAGGA 1885
Db 1622 TGCAGAACACTTGAACATGCTCATGACCTGCCACCACTAGATAAACGATCCGCGAGGA 1681
QY 1886 TTTCATATTTTCTCAAGCGGTATCCGCCCGGCTCTATCCGGCTGAAGATGTGCTCCA 1945
Db 1682 TTTCATATTTTCCCAAGCCGTATCCGCCCGGCTCTATCCGGCTGAAGATGTGCTCCA 1741
QY 1946 TGATATGGGTGATCGCGATGAACGATTCGGATTCGCAAGCAATGGCGCGTCGACGGCA 2005
Db 1742 TGATATTTGGGTGATCGCGATGAACGATTCGGATTCGCAAGCAATGGCGCGCTGGCA 1801
QY 2006 AGTGATTTCTCGAATTTGGCAGACTTGGCAGATGAAGATAAAGAAATTTGTAAGCTTC 2065
Db 1802 AGTGATTTCTCGAATTTGGCAGACTTGGCAGATGAAGATAAAGAAATTTGTAAGCTTC 1861
QY 2066 TGAAGATGGCAAGATAACGATAATTTCCGATTAAGCGCTACATCTCCAAATACACTAT 2125
Db 1862 TGAAGATAGTGCAGATAACGAACTTCCGTATCAAAACGCTACATCTCCAAATACACTAT 1921
QY 2126 CAACCCGCTTTGACCCACGCGTGAAGTATATCGGCTCTGTGGAAGAGGCAAGAT 2185
Db 1922 TAACCCGCTCTTAACCATGGGGTAAGCGATATATCGGCTCTGTGGAAGAGGCAAAAT 1981
QY 2186 CGCCGACTTGTGTGTGGAAATCCTGCTTTTGGCGTAAAAACCCAAATTCGTATCAA 2245
Db 1982 CGCTGATTTGGTGTGGAAATCCTGCTTTTGGTGTGAAACCTAAGATTTGTATCAA 2041
QY 2246 AGGCGGTATGGTGTCTTCTCTGAAAATGGGCGATTTTAAACGGCTCTGTGCCACTCCCA 2305
Db 2042 AGGCGGTATGGTGTCTTCTCTGAAAATGGGCGACTTCCAAACGCTCCGTGCTTACCTCA 2101
QY 2306 ACCGTTTATTTACCGCAAAATGTTTGGGCATCAGCGCAAGCGGAAATTTGACACGAT 2365
Db 2102 GCGGTTTATTTACCGCAAAATGTTTGGGCATCAGCGCAAGCGGAAATTTGACACGAT 2161
QY 2366 CACTTTTGTTCCAAAGTCCGCTATGAAAATGGGCTGAAAGAAAAGCTGGGGCTTAGAGCG 2425
Db 2162 CACTTTTGTTCCAAAGTCCGCTATGAAAATGGGCTGAAAGAAAAGCTTAGAGCGG 2221
QY 2426 CCAAGTTCTACCGTCAAAAATCTCCGCTAAACATCAACAAAGAAAGCTTCAAGTTCAACGA 2485
Db 2222 CAAGGTGTACCCGTGAAAATCTCCGCAACATCACTAAGAAAGACTTCAAAATTCAAACAA 2281
QY 2486 CAAAAACGGCAAAATCAACCGTCGATCCGAAACCTTCGAGGTCTTTGTAGATGGCAACT 2545
Db 2282 CAAGACGGCGCATATCACTGTGATCTTCAAAACCTTCGAGGTCTTTGTAGATGGCAACT 2341
QY 2546 CTGCACCTCTAAACCCACCTCGCAAGTGCCTCTAGCCAGCGCTACACTTCTTCTTAGGC 2605

| | | | |
|------------|--|---|-----------------|
| Db | 1323 | AGAAGCGGGCGGCGATTGGCTTTAAATTGCAATGAAGACTGGGGCACAAACACAAAGTGGCAT | 1382 |
| Qy | 1586 | CGATCACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATAC | 1645 |
| Db | 1383 | CGATCACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATAC | 1442 |
| Qy | 1646 | AGTCAATAGCGCAGGTTATGTAGATGACACCTTAATGCAATGAACGGGCGCGCATCCA | 1705 |
| Db | 1443 | GGTCAATAGCGCAGGTTATGTAGATGACACCTTAATGCAATGAACGGGCGCGCATCCA | 1502 |
| Qy | 1706 | TGCTACCACTTGAAGGAGGGGTGGAGGACACTGACCTGATGTTATCAACATGGCAGG | 1765 |
| Db | 1503 | TGCTACCACTTGAAGGAGGGGTGGAGGACACTGACCTGATGTTATCAACATGGCAGG | 1562 |
| Qy | 1766 | CGAGCTCAATATTCTACCTCTCCACACCCACATATTCCCTATACCAATTAATACGGT | 1825 |
| Db | 1563 | CGAGCTCAATATTCTACCTCTCCACACCCACATATTCCCTATACCAATTAATACGGT | 1622 |
| Qy | 1826 | TGCAGAACACTTAGACATGCTCATGACATGCCACCTAGACAAACGCATCCGCGAGGA | 1885 |
| Db | 1623 | TGCAGAACACTTAGACATGCTCATGACCTGCCACCTAGATAAGGCAATCCGCGAGGA | 1682 |
| Qy | 1886 | TTTACAAATTTCTCAAGCCGATATCGCCCGCGCTCTATCGCGCTGAAGATGTCTCCA | 1945 |
| Db | 1683 | TTTACAAATTTCTCAAGCCGATATCGCCCGCGATCTATTTCGCGCTGAGGATGTCTCCA | 1742 |
| Qy | 1946 | TGATATGGGTGATCGGATGACAAAGCTCGGATTCGCAAGCAATGGGGCGTCAGCGGA | 2005 |
| Db | 1743 | TGATATGGGTGATCGGATGACTAGCTCCGATTCGCAAGCAATGGGGCGTCGCGGA | 1802 |
| Qy | 2006 | AGTGATTCCTCGAACTTGGCAGACTCGGATTAAGATAAATAAAGAAATTCGTTAAGCTTCC | 2065 |
| Db | 1803 | AGTGATTCCTAGAACTTGGCAACTGCAGATAGATATAAATAAAGAAATTCGTTAAGCTTCC | 1862 |
| Qy | 2066 | TGAAGATGGCAAGATAACGATTAATTCGCACTTAAGCGCTACATCTCCAAATACACTAT | 2125 |
| Db | 1863 | TGAAGATGGTGAGATAAACGACAACTTCGCGCATCAAAAGCTATATCTCCAAATACACCAT | 1922 |
| Qy | 2126 | CAACCCCGCTTACCCACCGCGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGAAGAT | 2185 |
| Db | 1923 | TAATCCCGCTTACCCACCATGGCGTGAAGCGAGTATATCGGCTCTGTGGAAGAGGCGAAGAT | 1982 |
| Qy | 2186 | CGCCGACTTGGTGTGGAACTCTGCTTTTGGGTAAAGTAAACCCAAATCTGTATCAA | 2245 |
| Db | 1983 | CGCCGACTTGGTGTGGAACTCTGCTTTTGGGTAAAGTAAACCCAAATCTGTATCAA | 2042 |
| Qy | 2246 | AGCGGTATGGTGTCTTCTCTGAAATGGCGGATTCCTAACGCGTCTGTGCCCACTCCCCA | 2305 |
| Db | 2043 | AGGTGCATGGTGTCTTCTCTGAAATGGCGGATTCCTAACGCGTCTGTGCCCACTCCCA | 2102 |
| Qy | 2306 | ACCGGTTTATACCGGAAATGTTGGGCAATCAAGGCAAGGCGAAATTTGACACCGAGCAT | 2365 |
| Db | 2103 | GCCGGTTTATACCGGAAATGTTGGGCAATCAAGGCAAGGCGAAATTTGACACCGAGCAT | 2162 |
| Qy | 2366 | CACCTTTGTTCCAAAGTCG | 2385 |
| Db | 2163 | CACCTTTGTTCTCAAGCG | 2182 |
| RESULT 10 | | | |
| AX356689 | | | |
| LOCUS | AX356689 | 2183 bp | DNA |
| DEFINITION | Sequence 7 from Patent EP1176192. | linear | PAT 13-FEB-2002 |
| ACCESSION | AX356689 | | |
| VERSION | AX356689.1 | GI:18674026 | |
| KEYWORDS | Helicobacter felis | | |
| SOURCE | Helicobacter felis | | |
| ORGANISM | Helicobacter felis | | |
| | Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; | | |
| | Helicobacteraceae; Helicobacter. | | |
| REFERENCE | 1 | | |
| AUTHORS | Kusters, J. G. and Cattoli, G. | | |
| TITLE | Helicobacter felis vaccine | | |

| | | | |
|----------|---|---|-----|
| JOURNAL | Patent: BP 1176192-A 7 30-JAN-2002; | | |
| FEATURES | Akzo Nobel N.V. (NL) | | |
| source | Location/Qualifiers | | |
| | 1. .2183 | | |
| | /organism="Helicobacter felis" | | |
| | /mol_type="unassigned DNA" | | |
| | /db_xref="taxon:214" | | |
| CDS | 3. .683 | | |
| | /note="unnamed protein product" | | |
| | /codon_start=1 | | |
| | /transl_table=11 | | |
| | /protein_id="CAD23259.1" | | |
| | /db_xref="GI:18674027" | | |
| | /translation="MKLTPREQEFLLYAGEVARKKAELKLNQPEAIAYISAHIM | | |
| | DEARRGKKTVAELMEBCWFLFKDDEVPVGMVMDLQVEATFDDGTLKLVTVNWDIEP | | |
| | DEHFKAGEVFGCDKDIELANGKEVTELEVNKGPKSLHVGSHFHFPTNKAJFKDRE | | |
| | KAYGKRLDIFSGNTRLRIGAGTRKVLQPLGGSSKKVIGMNLVNNIADERHKKHAKLDK | | |
| | AKSHGFIK" | | |
| CDS | 694. .>2181 | | |
| | /note="unnamed protein product" | | |
| | /codon_start=1 | | |
| | /transl_table=11 | | |
| | /protein_id="CAD23260.1" | | |
| | /db_xref="GI:18674028" | | |
| | /translation="MKMKQEVYNTYPTTGDVKVRLGDTDLWAEVEHDDYTTVGEELKF | | |
| | GAGKTIREGMQSGSPDENTLDLITNAMIIDYTIYKADIGIKNGALHIGIKAGNKD | | |
| | MODGVSMVVGQTEALAGEMLITAGGIDSHFLSPQOFKXKALANGVITMPGGGT | | |
| | GPVDTGNATITPGKNLHMLRAAEYSMNVGFLGKGNSSSKQOLVQVEAGAIGFK | | |
| | LHEDWGTPPSAIDHCLSVADVEDVQVCITHTYINTVAEHLMDMLCHLDRIRREDIQFS | | |
| | AGGSDPVDITMAGELNLPSTPTPTINTVAEHLMDMLCHLDRIRREDIQFS | | |
| | QSIRPGSIAAEDVLHDIGVATMSDSQAMGRAGEVIRPTWQADKNKKEFGKLPEP | | |
| | GANDNPRIKRYISKYITNPALTGVSEYIGSVEEGKIADLVVWNPAPFGVKPKIVIK | | |
| | GMVVPSEMGSNDSNVPTPQVPVYREMFHGHGKAKFDITSITFRVSS" | | |
| ORIGIN | | | |
| | Query Match 67.6%; Score 1948; DB 6; Length 2183; | | |
| | Best Local Similarity 93.3%; Pred. No. 0; | | |
| | Matches 2035; Conservative 0; Mismatches 145; Indels 0; Gaps 0; | | |
| Qy | 206 | GTGAACTCAACCCAAAGAGCAAGAAAGTCTTGTATATATATATGCGGGCGAAGTGGCT | 265 |
| Db | 3 | GTGAACTCAACCCAAAGAGCAAGAAAGTCTTGTATATATATATGCGGGCGAAGTGGCT | 62 |
| Qy | 266 | AGAAAGCGCAAGCAGAGGCTTAAAGCTCAACCCGAAAGCCATTGCTTACATTAAGT | 325 |
| Db | 63 | AGAAAGCGCAAGCAGAGGCTTAAAGCTCAACCCGAAAGCCATTGCTTACATTAAGT | 122 |
| Qy | 326 | GCCCATATTATGGAACGAGCGCGCTGGAAAAAACCCTTGCCAGCTTATGGAAGAG | 385 |
| Db | 123 | GCCCATATTATGGAACGAGCGCGCTGGAAAAAACCCTTGCTGAACCTTATGGAAGAA | 182 |
| Qy | 386 | TGCATGACATTTTGAAGAAAGATGAAGTAATGCCCGGGGTGGGTAAATATGGTTCCCGAT | 445 |
| Db | 183 | TGTATGACATTTTGAAGAAAGATGAGTGATGCGCGGTGGGGAAATATGGTCCCTGAT | 242 |
| Qy | 446 | CTAGGTGTAGACCCACTTTCCCTGATGGTAGAACTTGTAACTGTGAATTGGCCCATC | 505 |
| Db | 243 | TTGGGCGGTAGAAGCCACTTTCCCGATGGCAACCAACTCGTAACCGTGAATTGGCCCAT | 302 |
| Qy | 506 | GAACCATAGCAGCACTTCAAGCGGGCGAAGTGAATTTGGTTGCGATAAAGACATCGAG | 565 |
| Db | 303 | GAACCTGATGACACTTTAAAGCGCGTGAAGTGAATTTGGCTGTGATTAAGACATTCAG | 362 |
| Qy | 566 | CTCAATGACGCAAGAAAGTAAACCGAATTGAGGTACTTAATGAAGGCGCTAAATCTCTTG | 625 |
| Db | 363 | CTCAACGTGGTAAAGAAAGTTACCGAGCTTGAAGTTACCAACGAAGGACCTAAATCTCTTG | 422 |
| Qy | 626 | CATGTGGGTAGCCATTTCCACTTCTTTGAAGCTAACAGGCACTAAATTCGATCGTGA | 685 |
| Db | 423 | CATGTGGGTAGCCATTTCCACTTCTTTGAAGCTAACAGGCACTAAATTCGATCGGGA | 482 |
| Qy | 686 | AAAGCCTATGCAACCGCTAGATATTCCCTCTCGCAACACGCTACGCAATTGGGGCAGGA | 745 |

| | | | |
|----|------|--|------|
| Db | 483 | AAAGCCTATGGCAACGCCCTAGATATTCCTCTGGCAACACGCTACGCATTGGGGCAGGA | 542 |
| QY | 746 | CAAAACCCGCAAAAGTGCAGTTGATTCCTCTTGGTGGCAGTAAAAAAGTGATGGCATGAAC | 805 |
| Db | 543 | CAAAACCCGTAAGTGCAGTTAATCCCTCTTGGCGGTAGTAAAAAAGTGATGGCATGAAC | 602 |
| QY | 806 | GGGCTTGTGAATAACATCCGGATGAACGCCATAAAACATAAAGCGCTTGACAAAGGGCAAA | 865 |
| Db | 603 | GGGCTTGTGAATAATATTCGACGCAACGCCATAAAACAAAACACTAGACAAAGGCAAAA | 662 |
| QY | 866 | TCTCAGCGATTTATCAAGTAAGGAGACTCCCATGAAATGAAATAACAAAACAGAAATATGTAAA | 925 |
| Db | 663 | TCTCAGCGATTTATCAAGTAAGGAGACTCCCATGAAATGAAATAACAAAACAGAAATATGTAAA | 722 |
| QY | 926 | TACCTACGACCCACCAAGGGGATAAAGTGGCTTTAGGAGATACCGATCTTTGGGCAGA | 985 |
| Db | 723 | CACCTACGACCCACCAAGGGGATAAAGTGGCTTTAGGAGATACCGATCTTTGGGCAGA | 782 |
| QY | 986 | AGTAGAACATGACTATACCACTATGGCGAAGAACTTAAATTTGGCGCGGTAAAACTAT | 1045 |
| Db | 783 | AGTAGAACATGACTATACCACTATGGCGAAGAGCTCAAATTTGGCGCGGTAAAACTAT | 842 |
| QY | 1046 | CGGTGAGGGTATGGGTACAGCAATAGCCCTGATGAAACACCCCTAGATTTAGTCTAC | 1105 |
| Db | 843 | CGGTGAGGGTATGGGTACAGCAATAGCCCAAGATGAAACACCCCTAGATTTAGTCTAC | 902 |
| QY | 1106 | TAAACGCGATGATTTACGACTACACCCGGGATTTACAAAGCCGATTTGGGATTAATAACGG | 1165 |
| Db | 903 | CAACGCGATGATTTACGACTACACCCGGGATTTAAAGCCGATTTGGTATTAATAATGG | 962 |
| QY | 1166 | CAAAATCCATGGCATTTGGCAAGCAGGAACAAAGACATGCAAGATGGCGTAAAGCCCTCA | 1225 |
| Db | 963 | CAAAATCCATGGTATTTGGCAAGCGGGAAACAAAGACATGCAAGATGGCGTAAAGCCCTCA | 1022 |
| QY | 1226 | TATGTGCGTGGGTGGGCACAGAACACTAGCAGGGGAGGTATGATTTATACCGCTGG | 1285 |
| Db | 1023 | TATGTGCGTGGGTGGGCACAGAACACTAGCAGGGGAGGTATGATTTATACCGCTGG | 1082 |
| QY | 1286 | GGGAATCGATTCACACACCCACTTCCTTCTCCACAAACAAATCCCTACCGCTCTAGCCAA | 1345 |
| Db | 1083 | GGGGATCGATTCGACACCCACTTCCTTCTCCCAACAAATCCCTACCGCTCTAGCCAA | 1142 |
| QY | 1346 | TGGCGTTACACCATGTTTGGAGCGGCACAGGTCTCTGTAGATGGCACGAATCGCACTAC | 1405 |
| Db | 1143 | TGGTGTACACCATGTTTGGAGTGGCACAGGTCCGCTAGATGGCACGAATCGCACTAC | 1202 |
| QY | 1406 | TATCACTCGGGCAATGGAACTTGCACCGCATGTTGGCGGACGAGAGAGATTTCTAT | 1465 |
| Db | 1203 | CATCACTCGGGCAATGGAACTTGCACCGCATGTTGGCGGACGCTGAAGAGTATTTCTAT | 1262 |
| QY | 1466 | GAATGTGGCTTTTGGGCAAGGCAATAGCTCTAGCAACCAATCTGTAGAACAAAT | 1525 |
| Db | 1263 | GAATGTAGGCTTTTGGGCAAGGCAATAGTCTAGCAACCAATCTGTAGAACAAAT | 1322 |
| QY | 1526 | AGAAGCGGGCGCATTTGGTTTAAATTCATGAAGACTGGGGCACCAACACCAAGTGCAT | 1585 |
| Db | 1323 | AGBAGCGGGCGCATTTGGCTTTAAATTCATGAAGACTGGGGCACCAACCAAGTGCAT | 1382 |
| QY | 1586 | CGATCACTGCTTGGCGTGGCAGATGAATACGATGCAAGTGTGATTTGATCCACACCGATAC | 1645 |
| Db | 1383 | CGATCACTGCTTGGCGTGGCAGATGAATACGATGCAAGTGTGATTTGATCCACACCGATAC | 1442 |
| QY | 1646 | AGTCAATGAGCGAGTTATGATGACACCCCTAAATGCAATGAACGGGCGGCCATCCA | 1705 |
| Db | 1443 | GGTCAATGAGCGAGTTATGATGACACCCCTAAATGCAATGAACGGGCGGCCATCCA | 1502 |
| QY | 1706 | TGCTTACCACATTGAGGGCGGGTGGAGGACACTCACCTGATGTTATACCATGGCAGG | 1765 |
| Db | 1503 | TGCTTACCACATTGAGGGCGGGCGGAGGACACTCACCTGATGTTATACCATGGCAGG | 1562 |
| QY | 1766 | CGAGCTCAATATTCTACCTCTCCACACCCCACTATTCCCTATACCAATTAATACGGT | 1825 |
| Db | 1563 | CGAGCTCAATATTCTACCTCTCCACACCCCACTATTCCCTATACCAATTAATACGGT | 1622 |

| | | | |
|----|------|--|------|
| QY | 1826 | TGCAGAACACTTATAGACATGCTCATGATGCCACCACTTAGACAAACGCAATCCGCGAGGA | 1885 |
| Db | 1623 | TGCAGAACACTTATAGACATGCTCATGACCTGCCACCACTTAGATAAGCGCATCCGCGAGGA | 1682 |
| QY | 1886 | TTTACAAATTTTCTCAAAGCGGTATCCGCCCGGCTCTATCGCGCTGAAGATGCTTCCA | 1945 |
| Db | 1683 | TTTACAAATTTTCTCAAAGCGGTATCCGCCCGGCTCTATTTGCCGCTGAGGATGTGCTCCA | 1742 |
| QY | 1946 | TGATATGGGTGTGATCGCGATGACAAAGCTCGGATTCGCAAGCAATGGGGCTGCGAGGGA | 2005 |
| Db | 1743 | TGATATTTGGGTGTGATCGCGATGACAAAGCTCGGATTCGCAAGCAATGGGGCTGCGAGGGA | 1802 |
| QY | 2006 | AGTGATTTCTCGAACTTTGGCAGACTGCGGATAAGAAATAAAAAAGAAATTTGGTAAAGCTTCC | 2065 |
| Db | 1803 | AGTGATTTCTTAGAACTTTGGCAAACTGCAGATAAGATAAAAAAGAAATTTGGTAAAGCTTCC | 1862 |
| QY | 2066 | TGAAGATGGCAAGATAACGATTAATTTCCGCAATTAAGCGGTACATCTCCAATACACTAT | 2125 |
| Db | 1863 | TGAAGATGGTGCAGATAACGCAAACTTTCCGCAATTAAGCGGTATATCTCCAATACACTAT | 1922 |
| QY | 2126 | CAACCCCGCTTTTGACCCACCGCGTGAAGAGTATATCGGCTCTGTGGAAGAGGGCAAGAT | 2185 |
| Db | 1923 | TAATCCCGCTTTTGACCCATGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGGCAAGAT | 1982 |
| QY | 2186 | CGCCGACTTGGTGGTGGGAATCCTGCCTTTTTTGGCGTAAAAACCCAAATTCGTGATCAA | 2245 |
| Db | 1983 | CGCCGACTTGGTGGTGGGAATCCTGCCTTTTTTGGCGTGAACCTTAAGATTGTGATTAA | 2042 |
| QY | 2246 | AGCGGATAGGTGGTCTTCTCTGAAATGGCGGATTTCAAACGCTCTGTGCCCACTCCCA | 2305 |
| Db | 2043 | AGTGCGCATGGTGGTCTTCTCTGAAATGGGCGATTTCAAACGCTCTCGGTGCCACCGCTCA | 2102 |
| QY | 2306 | ACCGGTTTATTACCGGCAATGTTTGGGSCATCAGGCAAGCGAAATTTTGACACACGAGAT | 2365 |
| Db | 2103 | GCCGTTTATTACCGCAATGTTTGGGACACCGCAAGCGAAATTTTGACACACGAGAT | 2162 |
| QY | 2366 | CACCTTTGTTTCCAAAGTCG 2385 | |
| Db | 2163 | CACCTTTGTTCTCAAGCG 2182 | |

| | | | | | |
|------------|---|-------------|-----|--------|-----------------|
| RESULT 11 | AF330621 | 8406 bp | DNA | linear | BCT 13-MAY-2003 |
| LOCUS | AF330621 | | | | |
| DEFINITION | Helicobacter bizzozeronii tRNA ribosyltransferase gene, partial cds; GIB, complete cds; urease gene cluster, complete sequence and unknown gene. | | | | |
| ACCESSION | AF330621 | | | | |
| VERSION | AF330621.1 | GI:27462193 | | | |
| KEYWORDS | Helicobacter bizzozeronii | | | | |
| SOURCE | Helicobacter bizzozeronii | | | | |
| ORGANISM | Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales; Helicobacteraceae; Helicobacter. | | | | |
| REFERENCE | 1 (bases 1 to 8406) | | | | |
| AUTHORS | Zhu,J., Teng,C.H., Chang,C.F., Chang,C.D., Simpson,K.W., Wei,C., McDonough,P., McDonough,S. and Chang,Y.F. | | | | |
| TITLE | Cloning and characterization of a Helicobacter bizzozeronii urease gene cluster | | | | |
| JOURNAL | DNA Seq. 13 (6), 321-331 (2002) | | | | |
| PUBMED | 12652903 | | | | |
| REFERENCE | 2 (bases 1 to 8406) | | | | |
| AUTHORS | Zhu,J. and Chang,Y.F. | | | | |
| TITLE | Direct Submission | | | | |
| JOURNAL | Submitted (21-DEC-2000) Population Medicine and Diagnostic Science, College of Veterinary Medicine, Cornell University, Tower Road, Ithaca, NY 14853, USA | | | | |
| FEATURES | Location/Qualifiers | | | | |
| source | 1. .8406 | | | | |
| | /organism="Helicobacter bizzozeronii" | | | | |
| | /mol_type="genomic DNA" | | | | |
| | /db_xref="taxon:56877" | | | | |

HECUREASE
LOCUS 2664 bp DNA linear BCT 03-MAY-1995
DEFINITION Helicobacter heilmannii urease, complete cds's.
ACCESSION U00002
VERSION 1.25079.1 GI:793908
KEYWORDS urease.
SOURCE Candidatus Helicobacter heilmannii
ORGANISM Candidatus Helicobacter heilmannii
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Helicobacteraceae; Helicobacter.
REFERENCE 1 (bases 1 to 2664)
AUTHORS Solnick J.V., O'Rourke J., Lee A. and Tompkins L.S.
TITLE Molecular analysis of urease genes from a newly identified
uncultured species of Helicobacter
JOURNAL Infect. Immun. 62 (5), 1631-1638 (1994)
PUBMED 8168924
COMMENT On May 3, 1995 this sequence version replaced gi:529422.
Original source text: Helicobacter heilmannii (individual_isolate
2) DNA.

FEATURES
Location/Qualifiers
1..2664
 /organism="Candidatus Helicobacter heilmannii"
 /mol_type="genomic DNA"
 /isolate="2"
 /db_xref="taxon:35817"
 /lab_host="BALB/c"
 200..204
 211..915
 /codon_start=1
 /transl_table=11
 /product="urease"
 /protein_id="AAA65722.1"
 /db_xref="GI:529423"
 /translation="MKLTPEKDKMLHYAGELAKORAKGIGIKLNYTEAVALISAHVM
EEARAPGSVADLMQSGRTLLKADDVPWGAHMEHVGLIAGPPDCKLVTHPTVEA
GSKLPAGEVLNKGRIITLNAGKHAVQLKVKNKGRDPVQGVSHFHFVFNKLDFPRE
KAYGKKDLASGADVTFEPGESEETKIIYGFNALVDROHDHGKKALKLR
AKEHFGTINCQDNK"
 919..923
 RBS
 930..2636
 /codon_start=1
 /transl_table=11
 /product="urease"
 /protein_id="AAA65723.1"
 /db_xref="GI:529424"
 /translation="MKKISRKEVSMYGPTTGDKVRIGDTLDLILEVHDCCTTYGERIK
FGGKTIIRGMGTNSPSHSHELLDVITNALIVDTGYIKADIKIGKNGKHGIGKAGNK
DLQGVNRCVGPATEALAAEGLIVTAGGIDTHIFSPQQTIPAFAPSGITTMGIGG
TPGDATNTATTIPGRNLKEMLRABEAYEAMNLYLGKGVSEPALIDGLEAGAIGF
KIHHEDGSPSAINALNTADKYDVOVAIHDTLNEAGCVEDTLEAIAGRTIHTFTFE
GAGGHGAPVINMAGENFIPLASTNPPIPTTKNTEAHMDMLCHHLDKNIKEDVFA
DSRIPQTTAAEDKLDHMGIFSITSSDSQAMRGVGEIVTWTQADKNKKEFKRLPEE
KGNDFNRISKYISKYITPAITHGISSEYGVSEVGKYADLWMSAFPQIKENMLIK
GGFIALSQMGDNASIPTPOPVYIRFMFHGHGAKFDNTIFFVSQVAYENGIKHELGI
QRVLPVQXNCRNRTKKDLKFENDVTAHIEVNPNETYKVKVDGNEVTSHAADKLSLAQLYN
LP"

ORIGIN

Query Match 39.5%; Score 1138; DB 1; Length 2664;
Best Local Similarity 66.4%; Pred. No. 6.8e-217;
Matches 1709; Conservative 1; Mismatches 826; Indels 36; Gaps 4;

Qy 63 TTGTTAATRCTATTATAATTTTTTAATAACTTACTTATTATCATATATAATAATATTATT 122
||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 70 TTGTGAATAAGATTACAAGAATAATAAT--ATTATCTTTGGATAACAAATATTAAAC 127

Qy 123 ACTTATTATTAAGAAGTTAATAAAAAAGTAACCGAAATAGGCATATAATCCCAATGCCCTTTA 182
Db 128 AACCATTAGCCAARAAGATACTAANAATCGCTCGTTGATAGCTTCGCTATTCAATAACA 187

Qy 183 AAATTTAACACAAGGAGTAATAGGTGAAACTCACACCAAGAGCAAGAAAAGTCTTTGT 242
Db 188 CCATTTTATACGAGGAGTTTAGAGATGAACTGCACCCTAAGAGAGTTTGGATAGTTGATGC 247

QY 1292 CGATTACACACCCACTTCTCTTCTCCACAAATTTCCCTACCGCTCTAGCCAAATGGCGT 1351
DB 1328 CGACACCCACATCCACTTCTATTTCTCTCAACAAATCCCTACTGCTCTTGTAGCGGAAT 1387
QY 1352 TACACACCATGTTTGGAGCGGCACAGGTCTGTAGATGCGACGAATGCGACTACTATCAC 1411
DB 1388 CACCAACCATGATCGGGGGCGGACTGGACCTGCAGATGCGACCAACGCCACCAACCATCAC 1447
QY 1412 TCCGGGCAAAATGGAACTTTCACCGCATTTGGCGCGCAGCAGAGAGATATTTCTATGAATGT 1471
DB 1448 TCCCGCGCGTGGAACTTAAAGAAATGCTCGCGCTTCTGAAGATACGCCATGAACCT 1507
QY 1472 GGGCTTTTGGCAAGGCAATAGCTCTAGCAAAAAACAACTTTGTAGAAACAAGTAGAAGC 1531
DB 1508 TGGCTACCTTGGTAAGGGGAATGTGCTTTTGAACCTGTCTCATTTGACCGAGCTAGAAGC 1567
QY 1532 GGGCGCGATTGGTTTAAATTCATGAAGCTGGGGCACAACACCAAGTGGCGATCGATCA 1591
DB 1568 CGCGCGCATTTGGCTTTAAATCCACGAAGACTGGGGCAGCACACCTCGACCATCAACCA 1627
QY 1592 CTGCTTGAGCGTGGCAGATGAATACCATGTGCAAGTTTGTATCCACACGATACAGTCAA 1651
DB 1628 CGCTCTAAACATCGCTGAACAAATACATGTGCAAGTGGCGATCCACACGACACCTTGA 1687
QY 1652 TGAGGCGATTATGTAGATGACACCTTAATGAATGAACGGCGCGCCATCCATCCCTTA 1711
DB 1688 TGAAGCGGCTGTGTGAAGACACCTTAGAAGCATCGCTGGACGACCATCCACACCTT 1747
QY 1712 CCACATTGAGGAGCGGTGGAGGACATCACTGATGTTATCACCATGCGAGCGAGCT 1771
DB 1748 CCACACCGAAGTGTCTGGCGGGGACACGCTCTGACGATCAAAATGGCGGGCGAAT 1807
QY 1772 CAATATCTTACCTCTCCACACCCCACTATTTCCCTATACCATTAATACGTTTCGAGA 1831
DB 1808 TAACTCTCTCTGCTTACCAACCCCACTCCCTTTCCACCAAAAACAGAACCGGA 1867
QY 1832 ACACCTTAGACATGCTCATGACATGCCACCACTAGACAAACGATCCGGGAGGATTTACA 1891
DB 1868 ACACATGGACATGCTTATG--TGCCACCATTTGGATAAAACATCAAGAAGATGTGA 1924
QY 1892 ATTTCTCAAAGCGTATCGGCCCGGCTTATCGCGGTGAAGATGTGCTCCATGATAT 1951
DB 1925 ATTTGCGGACTCACGATATCGGCCCAACCAATTTGGCGGGAAGATAAATCCACGACAT 1984
QY 1952 GGGTGTATCGGATCAGACGCTCGATTGCGAAGCAATGGGCGTGCAGGCGCAAGTAT 2011
DB 1985 GGGGATTTTCTCATCACTAGCTCCGACTCTCAAGCCATGGGTGCGGTGGGCGAAGTAT 2044
QY 2012 TCTCGAACTTGGCAGACTCGGATGAAGATAAAAGAAATTTGGTAAAGCTTCTCTGAAGA 2071
DB 2045 CACCGGCACTTGGCAACAGCGGACAAACAAAGAAATTTGGCGGCTTGTCTGAAGA 2104
QY 2072 TGGCAAGATAACGATAATTTCCGCAATTAAGCGCTACATCTCAAAATACATATCAACCC 2131
DB 2105 AAAAGCGGACACGACAACTTCGGCATCAACGCTATATTTCCAAATACACCATCAACCC 2164
QY 2132 CGCTTTGACCCACCGGCTGAGCAGTATATCGGCTCTGTGGAGAGGGGCAAGATCGCGGA 2191
DB 2165 TGGCATCACACAGCGCATTTCTGAATATGTGCGCTCTGTGAAGATGGGTAATACGCTGA 2224
QY 2192 CTGTGGTGTGTGAATCTCGCTTTTGGCGTAAACCCCAAAATCGTGTATCAAAAGCGG 2251
DB 2225 CTGTGTGCTTTGGAGCGCTCGCTTTTGGCATTAACCCCAACATGATCATCAAGCGG 2284
QY 2252 TATGTGTGCTTCTCTGAAATGGGCGATTTCAACGCGTCTGTGCCCATCTCCCAACCGGT 2311
DB 2285 TTTTCATTGCGCTTTCTCAAAATGGCGATGCAACGCTTCTATCCCACTCCCCAACCGGT 2344
QY 2312 TTATTAACCGGAATGTTTGGGATACAGCGCAAGGGGAAATTTGACACCAAGATCACTTT 2371
DB 2345 GTACTACCGGAAATGTTTGGCGCACCAACCGGCAAGGCAAAATTTGACACCAACATCACTTT 2404
QY 2372 TGTTCCTCAAAGTCGCTATGAAATGGCGTGAAGAAAAGCTGGGCTTAGAGCGCCAAGT 2431

DB 2405 CGTGTCTCAAGTGGCTTATGAAAACGGCATTAACACACGAGTTGGCTTGCAAAAGATTGT 2464
QY 2432 TCTACCGGTCAAAAACATGCGTAACATCACCAAGAAAGACTTCAAGTTTCAACGACAAAAC 2491
DB 2465 GTTGTCTGTGAAAACATGCGCGCAACATCACCAAAAAGACCTTAAGTTCAACGATGTAC 2524
QY 2492 GGCAAAAATCACCGTCGATCCGAAAACCTTCGAGGTCTTTGTAGATGCAAACTCTGCAC 2551
DB 2525 CGCACACATCGAAGTCAACCTTGAACCTCAAAAGTGAAGTGGATGGCAACGAAGTTAC 2584
QY 2552 CTCTAAACCCACCTCGCAAGTCCCTTAGCCAGCGGTACACATTTCTTCTTAG 2603
DB 2585 CTCCCATGCGGCTGACAAATTTGACCTTAGCACACTCTATAACCTATTCTTAG 2636

RESULT 13
AE001446/c

LOCUS Helicobacter pylori, strain J99 section 7 of 132 of the complete genome.
DEFINITION Helicobacter pylori, strain J99 section 7 of 132 of the complete genome.
ACCESSION AE001446
VERSION AE001446.1 GI:4154573

KEYWORDS

SOURCE Helicobacter pylori J99
ORGANISM Helicobacter pylori J99
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Helicobacter.

REFERENCE

AUTHORS 1 (bases 1 to 12037)
Alm, R.A., Ling, L.S., Moir, D.T., King, B.L., Brown, E.D., Doig, P.C., Smith, D.R., Noonan, B., Guild, B.C., deJonge, B.L., Carmel, G., Tummino, P.J., Caruso, A., Uria-Nickelsen, M., Mills, D.M., Ives, C., Gibson, R., Merberg, D., Mills, S.D., Jiang, Q., Taylor, D.E., Vovis, G.F. and Trust, T.J.

TITLE

Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori

JOURNAL

PUBMED Nature 397 (6715), 176-180 (1999)
9923682

AUTHORS

TITLE 2 (bases 1 to 12037)
King, B.L., Alm, R.A. and Trust, T.J.

COMMENT

Address all correspondence to: hp@arch.us.astro.com or Richard A. Alm, Astra Research Center Boston, 128 Sidney Street, Cambridge, MA 02139, USA
Submitted (12-JAN-1999) Astra Research Center Boston, 128 Sidney Street, Cambridge, MA 02139, USA
Address all correspondence to: hp@arch.us.astro.com or Richard A. Alm, Astra Research Center Boston, 128 Sidney Street, Cambridge, MA, 02139. Lo-See L. Ling, Donald T. Moir, Douglas R. Smith, Braydon C. Guild, Gilles Carmel, Anthony Caruso, Debra M. Mills, Rene Gibson, and Gerald F. Vovis are with Genome Therapeutics Corporation, 100 Beaver Street, Waltham, MA, 02453. Qin Jiang and Diane E. Taylor are with the University of Alberta Department of Medical Microbiology and Immunology, Edmonton, Alberta, Canada, T6G 2H7 and the Canadian Bacterial Diseases Network. All other authors are with Astra Research Center Boston, 128 Sidney Street, Cambridge, MA, 02139. Putative identifications, sequence alignments, and name and sequence search capability are available at ARCB's World Wide Web site. (URL: http://www.astro-boston.com/hpylori).

FEATURES

source 1. 12037
/organism="Helicobacter pylori J99"
/mol_type="genomic DNA"
/strain="J99"
/db_xref="taxon:85963"
complement (143. 1852)
/gene="ureB"
/note="synonym: jhp0067"
complement (143. 1852)
/gene="ureB"
/note="similar to H. pylori 26695 gene HP0072"
/codon_start=1
/transl_table=11
/product="UREASE BETA SUBUNIT"
/protein_id="AAD05651.1"

| | | | | | |
|------------|---|--|------|----------|--|
| Db | 615 | CAGCGATCGCTCATGAGATTAGCGAGTATGATAGGTTCTGTAGAAAGTGGGCAAAAGTGCTG | 556 | FEATURES | Location/Qualifiers |
| QY | 2191 | ACTTGGTGGTGTGGAATCCTGCCTTTTTCGGCGTAAACCCAAAATCGTGATCAAAAGCG | 2250 | source | 1. .10903 |
| Db | 555 | ACTTGGTATTTGGAGTCCAGCATCTTTGGCGTGAACCCACATCATCATCAAGCGC | 496 | | /organism="Helicobacter pylori 26695" |
| QY | 2251 | GTAATGGTGGTCTTCTCTGAAATGGCGAATCTTAACGGCTCTGTGCCCACTCCCCAACCGG | 2310 | gene | /strain="26695" |
| Db | 495 | GAATTCATTGGTTAAGCAAAATGGCGGATGCGAAGCTTCTATCCCTACCCCAACACCGG | 436 | CDS | /db_xref="taxon:85962" |
| QY | 2311 | TTTATTACCCGGAATGTTTGGGCATCAGCGCAGGCGAAATTTGACACAGATCACTT | 2370 | | /complement (127. .1836) |
| Db | 435 | TTTATTACAGAGAAATGTTCCGCTCACCATGGTAAGCTAATACGATGCAAAACATCACTT | 376 | | /gene="HP0072" |
| QY | 2371 | TTGTTTCCAAAGTCGCTATGAAATGGCGTGAAGAAAGCTGGCTTAGAGCGCCAAG | 2430 | | /complement (127. .1836) |
| Db | 375 | TTGTGTCCTAAGCGCTTATGACAAAGGCAATTAAGAAGATTTAGGCTTGAAGACAAG | 316 | | /gene="HP0072" |
| QY | 2431 | TTCTACCGGTCAAAAACCTGCCGTAAACATCAACCAAGAAGACTTCAAGTTCAACGACAAA | 2490 | | /note="similar to GB:M60398 GB:X17079 SP:P14917 PID:149011 |
| Db | 315 | TGTTGCCGTTAAAAATTGCAGAAACATCATCTAAAAAAGACATGCAATTCATGACACTA | 256 | | GB:AE000511 percent identity: 100.00; identified by sequence similarity; putative" |
| QY | 2491 | CGGCAAAATCACGTCGATCCGAAACCTTCGAGGTCTTTGTAGATGGCAAACTCTGCA | 2550 | | /codon_start=1 |
| Db | 255 | CCGTCACATTTGAAGTCAATCTCTGAAACTTACCATGTGTTTCGTGGATGCGAAAGAGTCA | 196 | | /transl_table=11 |
| QY | 2551 | CTCTAAACCACTCGCAAGTCCTCTAGCCAGCGCTACACTTTCTCTTAGG | 2604 | | /product="urease beta subunit (urea amidohydrolase) |
| Db | 195 | CTTCTAAACCAAGCTAATAAGTGAAGTGGCTCAACTCTTAGCATTTTCTTAGG | 142 | gene | (ureB)" |
| RESULT 14 | | | | CDS | /protein_id="AAD07143.1" |
| AE000529/c | | | | | /db_xref="GI:2313153" |
| LOCUS | AE000529 | 10903 bp | DNA | linear | /translation="MKKISRKEYVSMYGTTTGDKVRLGDTDLIAEVEHDDYTIYGEELK |
| DEFINITION | Helicobacter pylori 26695 section 7 of 134 | | | | PGGKTLREGMSQNNPSKEBLDIITNALIVDTGTIYKADIGIKDKIAGIKGKGNK |
| ACCESSION | AE000529 | AE000511 | | | DMODGVKNLSVGPATEALAGEGLIVTAGGIDTHIFISPOOIPTAFAGSVTTMIGGG |
| VERSION | AE000529.1 | GI:2313152 | | | TGPADGTTNATITPGRNLKMWLRAAEYSNNLGFSLAKGNASNDASLADQIEAGAIGF |
| KEYWORDS | Helicobacter pylori 26695 | | | | KIHEDWGTTPSAINHALDVADKYDQVAIHDTTLNEAGCEDTMAAIAKRTWHITFHE |
| SOURCE | Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Helicobacter. | | | | GAGGCHAPDIIKVAEHNILPASTNPTIPTVTNTEAEHMDMLVCHHLDKSIKEDYQF |
| REFERENCE | 1 (bases 1 to 10903) | | | | ADSRIPOTIAAEDTLHDMGIFSISSDSQAMRGVGEVITRTQWADKNKGFRLKE |
| AUTHORS | Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C. | | | | EKGDNFRIKRYLSKYTINPAIAHGISSEYGVSVGVADLVLSWPAFGVKPNMII |
| | The complete genome sequence of the gastric pathogen Helicobacter pylori | | | | KGGFIALSQMGDANASIPTQPVYVREMFHHGKAKYDANITFVSOAAVDKGIKEELG |
| JOURNAL | Nature 388 (6642), 539-547 (1997) | | | | LERQVLPVKNCRNITKDMQNDTTHAIEVNPETHVTFVDGKEVTSKPANKVSLAQLF |
| PUBMED | 9252185 | | | | SIF" |
| AUTHORS | 2 (bases 1 to 10903) | | | | complement (1840. .2556) |
| | Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C. | | | | /gene="HP0073" |
| | The complete genome sequence of the gastric pathogen Helicobacter pylori | | | | complement (1840. .2556) |
| | Nature 388 (6642), 539-547 (1997) | | | | /gene="HP0073" |
| | 2 (bases 1 to 10903) | | | | /note="similar to SP:P14916 PID:149010 PID:43634 |
| | Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C. | | | | GB:AE000511 PID:1567116 percent identity: 100.00; identified by sequence similarity; putative" |
| | The complete genome sequence of the gastric pathogen Helicobacter pylori | | | | /codon_start=1 |
| | Nature 388 (6642), 539-547 (1997) | | | | /transl_table=11 |
| | 2 (bases 1 to 10903) | | | | /product="urease alpha subunit (ureA) (urea amidohydrolase)" |
| | Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C. | | | | /protein_id="AAD07144.1" |
| | The complete genome sequence of the gastric pathogen Helicobacter pylori | | | | /db_xref="GI:2313154" |
| | Nature 388 (6642), 539-547 (1997) | | | | /translation="MKLTPKELDKLMLHYAGELAKKRKEKGIKLVNVEAVALISAHIM |
| | 2 (bases 1 to 10903) | | | | BEARAGKTAELMQEGRTLLKPDMDGVASMIHEVGIEAMFPDGTKLVTVHTPTIEA |
| | Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C. | | | | NGKLVPELFLKNEEDITINEGKASVSKVKNVGDPRVQIGSHFHFVNRCLDFDREK |
| | The complete genome sequence of the gastric pathogen Helicobacter pylori | | | | TFGRKLDIASGTAVRPFEGEKSVELIDIGNRRIRTFGNALVDROADNESKKIALHRA |
| | Nature 388 (6642), 539-547 (1997) | | | | KERGFHGAKSDDNYVTIKF" |
| | 2 (bases 1 to 10903) | | | | complement (2809. .2881) |
| | Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C. | | | | /gene="tRNA-Val-2" |
| | The complete genome sequence of the gastric pathogen Helicobacter pylori | | | | complement (2809. .2881) |
| | Nature 388 (6642), 539-547 (1997) | | | | /gene="tRNA-Val-2" |
| | 2 (bases 1 to 10903) | | | | /product="tRNA-Val" |
| | Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C. | | | | /note="codon recognized: GUC" |
| | The complete genome sequence of the gastric pathogen Helicobacter pylori | | | | complement (2902. .3375) |
| | Nature 388 (6642), 539-547 (1997) | | | | /gene="HP0074" |
| | 2 (bases 1 to 10903) | | | | complement (2902. .3375) |
| | Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C. | | | | /gene="HP0074" |
| | The complete genome sequence of the gastric pathogen Helicobacter pylori | | | | /note="similar to GB:M60398 SP:P25178 PID:149009 |
| | Nature 388 (6642), 539-547 (1997) | | | | GB:AE000511 PID:2313155 percent identity: 100.00; identified by sequence similarity; putative" |
| | 2 (bases 1 to 10903) | | | | /codon_start=1 |
| | Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C. | | | | /transl_table=11 |
| | The complete genome sequence of the gastric pathogen Helicobacter pylori | | | | /product="lipoprotein signal peptidase (lppA)" |
| | Nature 388 (6642), 539-547 (1997) | | | | /protein_id="AAD07145.1" |
| | 2 (bases 1 to 10903) | | | | /db_xref="GI:2313155" |
| | Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C. | | | | /translation="MLKTTKSLVFMVGVFFLIFGVDOAIKYAILEGFRYSIMDIV |
| | The complete genome sequence of the gastric pathogen Helicobacter pylori | | | | LVFNKGVAFSLFLEGLKYLQIILLGLFIFLMRQRELFAKHAIEFGMVGAVGASIN |
| | Nature 388 (6642), 539-547 (1997) | | | | VLDRFVHGVDVYVYHYHGFDFFAIFNFDADMVMDVGVGLLLKQFFFFKQNKIKA" |
| | 2 (bases 1 to 10903) | | | | complement (3369. .4706) |
| | Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C. | | | | /gene="HP0075" |
| | The complete genome sequence of the gastric pathogen Helicobacter pylori | | | | complement (3369. .4706) |
| | Nature 388 (6642), 539-547 (1997) | | | | /gene="HP0075" |
| | 2 (bases 1 to 10903) | | | | |
| | Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C. | | | | |
| | The complete genome sequence of the gastric pathogen Helicobacter pylori | | | | |
| | Nature 388 (6642), 539-547 (1997) | | | | |
| | 2 (bases 1 to 10903) | | | | |
| | Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C. | | | | |
| | The complete genome sequence of the gastric pathogen Helicobacter pylori | | | | |
| | Nature 388 (6642), 539-547 (1997) | | | | |
| | 2 (bases 1 to 10903) | | | | |
| | Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C. | | | | |
| | The complete genome sequence of the gastric pathogen Helicobacter pylori | | | | |
| | Nature 388 (6642), 539-547 (1997) | | | | |
| | 2 (bases 1 to 10903) | | | | |
| | Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C. | | | | |
| | The complete genome sequence of the gastric pathogen Helicobacter pylori | | | | |
| | Nature 388 (6642), 539-547 (1997) | | | | |
| | 2 (bases 1 to 10903) | | | | |
| | Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C. | | | | |
| | The complete genome sequence of the gastric pathogen Helicobacter pylori | | | | |
| | Nature 388 (6642), 539-547 (1997) | | | | |
| | 2 (bases 1 to 10903) | | | | |
| | Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C. | | | | |
| | The complete genome sequence of the gastric pathogen Helicobacter pylori | | | | |
| | Nature 388 (6642), 539-547 (1997) | | | | |
| | 2 (bases 1 to 10903) | | | | |
| | Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C. | | | | |
| | The complete genome sequence of the gastric pathogen Helicobacter pylori | | | | |
| | Nature 388 (6642), 539-547 (1997) | | | | |
| | 2 (bases 1 to 10903) | | | | |
| | Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C. | | | | |
| | The complete genome sequence of the gastric pathogen Helicobacter pylori | | | | |
| | Nature 388 (6642), 539-547 (1997) | | | | |
| | 2 (bases 1 to 10903) | | | | |
| | Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C. | | | | |
| | The complete genome sequence of the gastric pathogen Helicobacter pylori | | | | |
| | Nature 388 (6642), 539-547 (1997) | | | | |
| | 2 (bases 1 to 10903) | | | | |
| | Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C. | | | | |
| | The complete genome sequence of the gastric pathogen Helicobacter pylori | | | | |
| | Nature 388 (6642), 539-547 (1997) | | | | |
| | 2 (bases 1 to 10903) | | | | |
| | Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C. | | | | |
| | The complete genome sequence of the gastric pathogen Helicobacter pylori | | | | |
| | Nature 388 (6642), 539-547 (1997) | | | | |
| | 2 (bases 1 to 10903) | | | | |
| | Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C. | | | | |
| | The complete genome sequence of the gastric pathogen Helicobacter pylori | | | | |
| | Nature 388 (6642), 539-547 (1997) | | | | |
| | 2 (bases 1 to 10903) | | | | |
| | Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C. | | | | |
| | The complete genome sequence of the gastric pathogen Helicobacter pylori | | | | |
| | Nature 388 (6642), 539-547 (1997) | | | | |
| | 2 (bases 1 to 10903) | | | | |
| | Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C. | | | | |
| | The complete genome sequence of the gastric pathogen Helicobacter pylori | | | | |
| | Nature 388 (6642), 539-547 (1997) | | | | |
| | 2 (bases 1 to 10903) | | | | |
| | Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C. | | | | |
| | The complete genome sequence of the gastric pathogen Helicobacter pylori | | | | |
| | Nature 388 (6642), 539-547 (1997) | | | | |
| | 2 (bases 1 to 10903) | | | | |
| | Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C. | | | | |
| | The complete genome sequence of the gastric pathogen Helicobacter pylori | | | | |
| | Nature 388 (6642), 539-547 (1997) | | | | |
| | 2 (bases 1 to 10903) | | | | |
| | | | | | |


```
Db 1803 ATGTATGGCCCTACTACAGGGGATAAAGTGAGATTGGCGGATACAGACTTGATCGCTGAA 1744
Qy 987 GTAGAACATGACTATACCACTATGGCGAAGAACTTAAATTTGGCGGGTAAACATATC 1046
Db 1743 GTAGAACATGACTATACCACTTATGGCGAAGAGCTTAAATTCGGTGGGGTAAACCCGTG 1684
Qy 1047 CGTAGGGTATGGGTAGAGCAATAGCCCTGATGAAGAACCCCTAGATTTAGTCACTACT 1106
Db 1683 AGAGAAGGCATGAGCCCAATCCAAACAACCTAGCAAGAAGAATTTGGATCTAATCATCACT 1624
Qy 1107 AACCGCATGATTATCGACTACACCGGATTTACAAAGCCGACATTTGGGATTTAAAGCGG 1166
Db 1623 AACGCTTTAATCGTGGATATACCCGGTATTATTAAGCGGATATTGGTATTAAAGATGGC 1564
Qy 1167 AAAATCCATGGCATTTGCCAAGGAGGAAACAAGACATGCAAGATGGCGTAAGCCCTCAT 1226
Db 1563 AAAATCGCTGGCATTTGGTAAGCGGTAAACAAGACATGCAAGATGGCGTTAAAGACAT 1504
Qy 1227 ATGGTCGTGGGTGGGCAAGAGCACTAGCAGGGGAAGGTATGATTATTTACCGCTGGG 1286
Db 1503 CTTAGCGTAGGTCCTGCTACTGAAGCCTTAGCCGGTGAAGGTTTGATCGTAACCTGCTGGT 1444
Qy 1287 GGAATCGATTACACACCCACTTCTTCTCCACAACAATCCCTACCGCTCTAGCCAT 1346
Db 1443 GGTATTGACACACATCCCACTTCATTTCAACCCCAACAAATCCCTACAGCTTTTGCAGC 1384
Qy 1347 GCGGTTACAAACATGTTTGGAGGGCGGACAGGTCCTGTAGATGCGACGAATGCCACTACT 1406
Db 1383 GGTGTAAACAACATGATTGGTGGCGGAACCTGGTCTCTGTATGCACTAATTCGCACTACT 1324
Qy 1407 ATCACTCCGGGCAAAATGGAACTTTGCACCGCATGTTGGCGGAGCAGAGAGTATTTCTATG 1466
Db 1323 ATCACTCCAGGCGAAGAAATTTAAATGGATGCTCAGAGCGGCTGAAGAATAATTTCTATG 1264
Qy 1467 AATGTGGGCTTTTGGGCAAGGCAATAGCTCTTAGCAAAAAACAACCTTTGTAGAACAGTA 1526
Db 1263 AACTTAGGTTTCTTGGCTAAGGTAACGCTTTTAAACGACGCGAGCTTAGCGCATCAAAAT 1204
Qy 1527 GAAGCGGGCGGATGGTGTAAATTCATGAGACTGGGGCAACACCAAGTGGCGATC 1586
Db 1203 GAAGCTGGTCGATTGGCTTTAAATCCAGAAAGACTGGGGCACACTCTCTTCTGCAATC 1144
Qy 1587 GATCACTGCTTGACGCTGGCAGATGAATAGCATGTGCAAGTTTGTATCCACACCGATACA 1646
Db 1143 AATCATCGCTTAGATGTTGCAGACAAATACGATGTGCAAGTCGCTATCCACAGACACT 1084
Qy 1647 GTCAATGAGCAGGTTATGTAGATGACACCCCTAAATGCAATGAACGCGCGCGCCATCCAT 1706
Db 1083 TTGAATGAAGCCGCTTGGTGGGAAGACACTATGGCAGCTATTGGCGGACGCACTATGCAC 1024
Qy 1707 GCCTACCAATTTAGGAGCGGGTGGAGGACACTCACTGATGTTATCACCATGGCAGGC 1766
Db 1023 ACTTTCCACATGAAGGTGTGGCGGGGACACGCTCTGATATATTAAAGTAGCTGGT 964
Qy 1767 GAGCTCAATATTTACCCCTCTCCACACACCCCACTATTCCTATACATTTAATACGGTT 1826
Db 963 GAACACAACTTCTCCGCTTCCACTAACCCCACTATCCCTTCTCATGTGAATACAGAA 904
Qy 1827 GCAGAACACTTAGACATGCTCATGACATGGCACCACTAGACAAACGCAATCCCGGAGGAT 1886
Db 903 GCAGAACACATGGACATGCTTATGTTGGTGCACCACTTGGATATAAGCAATTAAGAAGAT 844
Qy 1887 TTAACAATTTCTAAGCCGTATCCGCCCGGCTCTATCGCGGCTGAAGATGCTCCAT 1946
Db 843 GTTCAGTTCGCTGATTCAAGGATCCGCCCTCAAAACCATTTGGCGTGAAGACACTTTGCA 784
Qy 1947 GATATGGTGTGATCGCGATGACAGACTCGGATTCGCAAGCAATGGGCGGTGAGGCGAA 2006
Db 783 GACATGGGATTTTCTCAATCACAGCTCTGACTCTCAAGCTATGGGTCTGGTGGGTGAA 724
Qy 2007 GTGATTTCTCGAATTTGGCAGACTGCGGATGAAGATTAAGAAATTTGGTAAAGCTTCTCT 2066
Db 723 GTTATCACTAGAACTTGGCAACAGCTGACAAAAACAAAAAGAAATTTGGCCGCTTGAA 664
```

```
Qy 2067 GAAGATGSCAAAGATAACGATAAATTTCCGCATTAAGCGCTACATCTCCAAATACACTATC 2126
Db 663 GAAGAAAAGGCGATAACGACAACTTCAGATCAAAACGCTACTTGTCTAAATACACCAAT 604
Qy 2127 AACCCCGCTTTGACCCACGCGGTGAGAGTATATCGGCTCTGTGGAAGAGGCAAGATC 2186
Db 603 AACCCAGCGATCGCTCATGGGATTAGCGAGTATGAGGTTCTGTAGAGTGGGCAAGTG 544
Qy 2187 GCCGACTTGGTGGTGGAAATCCTGCTTTTGGCGTAAAAACCCAAAAATCGTGATCAA 2246
Db 543 GCTGACTTGGTATTTGTGAGTCCCGCATTTTGGCGTAAACCCCAACATGATCATCAA 484
Qy 2247 GCGGATATGGTGGTCTTCTCTGAATGGCGATTTAAACGCTGTGTGCGCATCTCCCAA 2306
Db 483 GCGGGTTCAITGGCTTGAATGGTGAACGCGTCTTATCCCTATCCCTACCCACAA 424
Qy 2307 CCGGTTTATACCGGAAATGTTGGGCTACCGCAAGCGGAAATTTGACACACGATC 2366
Db 423 CCAGTTTATACAGAGAAATGTTGCTCATCATGTTAAAGCCAAATACGATGCAAAATC 364
Qy 2367 ACTTTGTTTCCAAAGTCGCTTATGAAATGGCGTAAAGAAAAGCTGGGCTTAGAGCG 2426
Db 363 ACTTTGTTCTCAAGCGCTTATGACAAAGCATTTAAAGAGATTTAGGCTTGAAGA 304
Qy 2427 CAAGTTCTACCGGTCAAAAACTGCGGTAAACATCACCAAGAAAGACTTCAAGTTCAACG 2486
Db 303 CAAGTTTGGCGTAAAAAAATTTGCAGAAACATCACTAAAAAAGACATGCAATTTCAACG 244
Qy 2487 AAAACGGCAAAATCACCGTCGATCCGAAACCTTGGAGGCTTTGTAGATGCAAACTC 2546
Db 243 ACTACCGCTCACTTGAAGTCAATCTGAAACTTACCATGTCTGTGGATGGCAAGAA 184
Qy 2547 TGCACCTTAACCCACCTCGCACTGCTCTAGCCAGCGCTACACTTTCTTCTAGG 2604
Db 183 GTAACTTCTAAACCGCAATAAAGTGAGCTTGGCGCAACTCTTTAGCATTTTCTAGG 126

RESULT 15
AR054309 LOCUS AR054309 2735 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5837240.
ACCESSION AR054309
VERSION AR054309.1 GI:5979886
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2735)
AUTHORS Lee, C.K., Monath, T.P., Ackerman, S.K., Thomas, W.D., Soman, G.,
Kleanthous, H., Weltzin, R.A., Pappo, J., Ermak, T., Guirakhoo, F.,
Bhagat, H. and Suesman, I.
TITLE Multimeric, recombinant urease vaccine
JOURNAL Patent: US 5837240-A 1 17-NOV-1998;
FEATURES
source
1..2735
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 37.3%; Score 1075.2; DB 6; Length 2735;
Best Local Similarity 66.0%; Pred No. 2.4e-204;
Matches 1622; Conservative 0; Mismatches 798; Indels 38; Gaps 3;

Qy 182 AAAATTTTAAACAAGGAGTAAATAGGTGAACACTCACACCCAAAGACGAGGCTTAAAGCTCAACCAA 301
Db 78 ACACCTTTAAGAAATAGGAGAAATGAGATGAACCTCACCCCAAGAGTTAGATAAGTTGATG 137
Qy 242 TTATATTATGCGGGCAAGTGGCTAGAAAGCGCAAGCGCAAGCGGCTTAAAGCTCAACCAA 301
Db 138 CTCCACTACGCTGGAGAAATGCGCTAAAAAACGCAAGAAAGGCAATTAAGCTTAACTAT 197
Qy 302 CCCGAAGCCATTGCTTTACATTAGTCCCATATTATTATGGACGAAGCGCGCTGGAAAAA 361
```

| | | | | |
|----|------|---|------|--|
| Db | 198 | GTAGAAGCAGTACGCTTTGATTTAGTGCCTCATATTTATGGAAGACGAGAGCTGTTAAAG | 257 | |
| Qy | 362 | ACCGTTGCCAGCTTATGAAGAGTGCATGCACCTTTTGGAAAAAGATGAAGTAATGCC | 421 | |
| Db | 258 | ACTGCCGCTGAATTTGATGACAGAGGGCGCATCTTTTAAACCAAGATGATGTGATGGAT | 317 | |
| Qy | 422 | GGGGTGGGTAATATGGTTCCCGATCTAGGTGTAGAGCCACCTTTCTCTGATGGTACGAAA | 481 | |
| Db | 318 | GGCGTGGCAAGCATGATCCATGAAGTGGGTATTGAAGCGATGTTCTCTGATGGACATAA | 377 | |
| Qy | 482 | CTTGTAACTGTGAATTTGGGCCCATCGAACCAAGATGAGCACTTCAAGCGGGCGAAGTGA | 541 | |
| Db | 378 | CTCGTAACCGTGCATACCCCTTATTTAGGCCAATGGTAAATTAGTTCCTGGTGAGTTG--- | 434 | |
| Qy | 542 | TTTGGTTGCGATAAAGACATCGAGCTCAATGACAGGCAAGAAAGTAACCGAATCTTGAGGTT | 601 | |
| Db | 435 | TTCTTAAAAATGAAGACATCACTATCAACGAAGCAAAAAGCCGTTAGCGTGAAGTT | 494 | |
| Qy | 602 | ACTAATGAAGGGCTAAATCTTTGTCATGTGGGTAGCCATTTCCACCTTTCTTTGAAGCTAAC | 661 | |
| Db | 495 | AAAAATGTTGGGACAGACGGTTCAAAATCGGCTCACACTTCCATTTCTTTGAAGTGAAT | 554 | |
| Qy | 662 | AAGGCATTAATTCGATCGTGAAGAGCCTATGGCAAGCCCTAGATATTCCTCTGGC | 721 | |
| Db | 555 | AGATGCTTAGACTTTGACAGAGAAAAAATTTCCGTTAAACGCTTAGACATTTGCGAGCGG | 614 | |
| Qy | 722 | AACACGCTAGCATTGGGCGAGCACAAACCCGCAAGTGCGATTGATTCCTCTGTGGGC | 781 | |
| Db | 615 | ACAGCGTTAAGATTGAGCGCTGGCGAAGAAAAATCCGTGAAATTTGATTGACATTTGGCGGT | 674 | |
| Qy | 782 | AGTAAAAAAGTGAATGGCATGAACGGGCTTTGTAATAACATCGCGATGAACGCCATAAA | 841 | |
| Db | 675 | AACAGAAGAACTCTTGGATTTAAGCGATTGTTGATAGACAAGCAGACAGAAAGCAA | 734 | |
| Qy | 842 | CATAAAGCGCTTGCAAGGCGAAATCTCACGGAATTT----- | 877 | |
| Db | 735 | AAAAATGCTTTACACAGAGCTAAAGAGCGTGGTTTTCATGGCGCTAAAGCGATGACAAC | 794 | |
| Qy | 878 | -----ATCAAGTAAGGAGACTCCCATGAAA--ATGAAAAACAAGATATGTAAT | 926 | |
| Db | 795 | TATGTAAAAACAATTAAGGAGTAAGAAATGAAAAAGATTAGCAGAAAAAGATATGTTCT | 854 | |
| Qy | 927 | ACCTACGGACCCCAAGCGGATAAAGTCGCTTAGGAGATACCGATCTTTGGGCGAGAA | 986 | |
| Db | 855 | ATGATGCTCTACTACAGCGGATAAAGTGAGATTGGGCGATACAGACTTGATCGCTGAA | 914 | |
| Qy | 987 | GTAGAAATGACTATACCACTATGGCGAAGAACTTAAATTTGGCGGGTAAACCTATC | 1046 | |
| Db | 915 | GTAGAAATGACTATACCACTTATGGCGAAGAGCTTAAATTCGGTGGCGGTAAACCCCTA | 974 | |
| Qy | 1047 | CGTGAGGTATGGTTCAGACCAATAGCCCTGATGAAACACCCCTAGATTTAGTCACT | 1106 | |
| Db | 975 | AGAGAAGGCATGAGCCCAATCTAACACCCCTAGCAAGAGAGTTGGATTTAAATATCACT | 1034 | |
| Qy | 1107 | AACGCGATGATTTACGACTACCGGATTTTACAAAGCCGACATTTGGGATTTAAAAACGGC | 1166 | |
| Db | 1035 | AACGCTTTAATCGTGAATTTACCGGATTTTAAAGCGGATATTTGTTATTTAAAGATGGC | 1094 | |
| Qy | 1167 | AAAAATCCATGGCATTTGGCAGGCGAGGAAACAAGGACATGCAAGATGGCGTAAAGCCCTCAT | 1226 | |
| Db | 1095 | AAAAATCGCTGGCATTTGGTAAGGCGGTAAACAAGACATGCAAGATGGCGTTAAAAACAAT | 1154 | |
| Qy | 1227 | ATGGTCTGGGTGGGCAAGAGCACTAGAGGGGAAAGGTATGATTTATACCGCTGGG | 1286 | |
| Db | 1155 | CTTAGCGTAGGCTCTGCTACTGTAAGCCCTTAGCGGTTGAAGTTTGTATCGTAACCGCTGGT | 1214 | |
| Qy | 1287 | GGATCGATTTACACACCCACTTCTTCTCCACAAATTCCTACCGCTCTAGCCAT | 1346 | |
| Db | 1215 | GGTATTGACACACATCCCACTTCAATTTCAACCCCAACAAATCCCTACAGCTTTTGCAGC | 1274 | |
| Qy | 1347 | GGCGTTACAAACCAATGTTGGAGGCGGCACAGGTCCTGTAGATGGCAGGAATGGCACTACT | 1406 | |
| Db | 1275 | GGTGTAAACAACCATGATTGGTGTGGAAACCGGTCCTGTGTATGGCACTAAATGGCACTACT | 1334 | |
| Qy | 1407 | ATCACTCCGGCGAAATGGAACTTGCACCGCATGTTGGCGCAGCAGAGAGATTTCTATG | 1466 | |
| Db | 1335 | ATCACTCCAGCAGAAAGAAATTTAAATGGATGCTCAGAGCGGCTGAAGAATATTTCTATG | 1394 | |
| Qy | 1467 | AATGTGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAACCTTTGTAGAAACAAGTA | 1526 | |
| Db | 1395 | AAATTTAGGTTTCTTGGCTAAAGGTAAGGCTTCTAACGATGCGAGCTTAGCGCATCAAAAT | 1454 | |
| Qy | 1527 | GAAGCGGCGGATTTGTTTAAATTCATGAAGCACTGGGGCAACAACCAAGTCGGATC | 1586 | |
| Db | 1455 | GAAGCGGTCGATTTGGCTTTTAAATTTACGAAGACTGGGGACCACTCTCTTCTGCAATC | 1514 | |
| Qy | 1587 | GATCACTGCTTTAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCCGATACA | 1646 | |
| Db | 1515 | AATCATCGGTTAGATTTTGGGACAAATACGATGTGCAAGTCGCTATCCACACACACT | 1574 | |
| Qy | 1647 | GTCAATGAGGAGGTTATGTATGACACCCCTAAATGCAATGAACGGGCGGCCATPCCAT | 1706 | |
| Db | 1575 | TTGAATGAAGCGGTTGTGTAGAAAGACACTATGGCTGCTATTGCTGGACGCACTATGCAC | 1634 | |
| Qy | 1707 | GCCTACACACTTGAAGGCGGTTGGAGGACACTCACTGTATGTTATCACCATGCGAGC | 1766 | |
| Db | 1635 | ACTTTCCACACTGAAGCGCTTGGCGGCGGACACGCTCCTGATATTTATTAAGTAGCCGT | 1694 | |
| Qy | 1767 | GAGCTCAATATCTACCCCTCCACACCCCTACTATTCCTATPACCATTAATACGGTT | 1826 | |
| Db | 1695 | GAACACAACATTTCTCCGCTTCCACTAACCCGACCATCCCTTTCACCGTGAATACAGAA | 1754 | |
| Qy | 1827 | GCAGAACACTTGTATGCTCATGACATGTCACCACTAGCAACCACTTAGCAACCAACGATCCGCGAGT | 1886 | |
| Db | 1755 | GCAGACACATGGACATGCTTATGGTGTGCCACCACTTTGGATTAAGAAGAT | 1814 | |
| Qy | 1887 | TTACAATTTTCTCAAGCCGTTATCGCCCCGGCTCTATCGCGGCTGAAGATGTGTCCAT | 1946 | |
| Db | 1815 | GTTCAAGTTCGCTGATCAAGGATCCGCCCTCAAAACCATTTGGCGCTGAAGACACTTTGGCAT | 1874 | |
| Qy | 1947 | GATATGGTGTGATCGGATGACAAAGCTCGGATTCGCAAGCAATCGGGGCTCAGCGCAA | 2006 | |
| Db | 1875 | GACATGGGATTTTCTCAATCACCAGTTCTGACTCTCAAGCGATGGGCGGTGGGTGAA | 1934 | |
| Qy | 2007 | GTGATTTCTCGAACTTTGGCAGACTCGGATTAAGAAATAAAAAAGAAATTTGGTAAGCTTCCT | 2066 | |
| Db | 1935 | GTTATCACTAGAACTTTGGCAACAGCTGACAAAAACAAGAAAGAAATTTGGCGGCTTGA | 1994 | |
| Qy | 2067 | GAGATGGCAAGATAAGATAAATTTCCGCAATTAAGCGCTCATCTCOAAATACACTATC | 2126 | |
| Db | 1995 | GAAGAAAAAGCGGATTAACGACAACTTCAGGATCAAAACGCTACTTGTCTAAATACACCAAT | 2054 | |
| Qy | 2127 | AACCCCGCTTTGACCCAGCGCTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGCAAGATC | 2186 | |
| Db | 2055 | AACCCAGCGATCGCTCATGGGATTAGCGAGTATGTAGGTTCAAGTAGAAGTGGGCAAGTG | 2114 | |
| Qy | 2187 | GCCGACTTGTGGTGTGGAATCTCGCTTTTGGCGTAAAAACCCAAAAATTCGTGATCAAAA | 2246 | |
| Db | 2115 | GCTGACTTGTATTTGTGGAGTCCAGCATTTCTTTGGCGTGAACCCCAACATGATCATCAA | 2174 | |
| Qy | 2247 | GGCGGTATGTGGTCTTCTCTGAAATGGCGGATTTAAACGCGTCTGTGCCCACTCCCCAA | 2306 | |
| Db | 2175 | GGCGGATTCATTTGGCTTAAGCCAAATGGGCGATGCGAACGCTTCTATCTCCTACCCCAAA | 2234 | |
| Qy | 2307 | CCGCTTTATACCGGAAATGTTTGGGCGATCACGGCAAGGGGAAATTTGACACCGCATC | 2366 | |
| Db | 2235 | CCGCTTTATACAGAGAAATGTTCCGCTCATCGTAAAGCTTAAATACGATGCAACATC | 2294 | |
| Qy | 2367 | ACTTTTGTTCCTCAAGTCCCTATGAAATCGCGTGAAGAAAGCTGGCGCTTAGAGCGC | 2426 | |
| Db | 2295 | ACTTTTGTGCTCAAGCGGCTTATGACAAAGCATTTAAAGAAAGATTTAGGACTTTGAAGA | 2354 | |
| Qy | 2427 | CAAGTCTTACCGGTCAAAAACCTGCGGTAAACATCACCAAGAAAGACTTCAAGTTCAACGAC | 2486 | |
| Db | 2355 | CAAGTGTTCGCGTAAAAAATTTGCAGAAATATCACTAAAAAAGACATGCAATTTCAACGAC | 2414 | |

| | | | |
|----|------|--|------|
| Qy | 2487 | AAAAACGGCAAAATCACCGTCGATCCGAAACCTTCGAGGTCTTTGTAGATGGCAAACTC | 2546 |
| Db | 2415 | ACTACTGTCTCACATTGAAGTCATCCTGAAACTTACCATGTGTTTCGTGGATGGCAAGAA | 2474 |
| Qy | 2547 | TGCACCTCTAAACCCACCTCGCAAGTGCCTCTAGCCGAGCGCTACACTTTTCTTCTAGG | 2604 |
| Db | 2475 | GTAACCTTCTAAACCCAGCAATAAAGTGAGCTTGGCGCAACTCTTTAGCAATTTCTAGG | 2532 |

Search completed: November 29, 2005, 04:30:37
Job time : 14066 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 23:31:00 ; Search time 10210 Seconds
(without alignments)
13211.279 Million cell updates/sec

Title: US-09-904-994B-1
Perfect score: 2883
Sequence: 1 rgrgagatttccaractt.....aaaaagtagagccacagg 2883

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 2339354128 residues

Total number of hits satisfying chosen parameters: 82156643

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_est3:*
- 4: gb_hic:*
- 5: gb_est4:*
- 6: gb_est5:*
- 7: gb_est6:*
- 8: gb_est7:*
- 9: gb_gss1:*
- 10: gb_gss2:*
- 11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB | ID | Description |
|------------|-------|-------|--------|----|----------|--------------------|
| 1 | 327.4 | 11.4 | 958 | 8 | DR637438 | DR637438 EST102806 |
| 2 | 324.2 | 11.2 | 946 | 8 | DR637093 | DR637093 EST102771 |
| 3 | 310.2 | 10.8 | 1509 | 4 | CNSOABR8 | BX816461 Arabidops |
| 4 | 309 | 10.7 | 872 | 7 | CO122599 | GR_EB04F |
| 5 | 306.2 | 10.6 | 925 | 8 | DR633422 | DR633422 EST102404 |
| 6 | 302.2 | 10.5 | 748 | 8 | CK676965 | Ydd29h04. |
| 7 | 299.2 | 10.4 | 889 | 8 | DR637386 | DR637386 EST102801 |
| 8 | 296.4 | 10.3 | 796 | 8 | DR633939 | DR633939 EST102456 |
| 9 | 292 | 10.1 | 849 | 8 | DR635599 | EST102622 |
| 10 | 279.6 | 9.7 | 766 | 7 | CV195867 | CV195867 CGF100344 |
| 11 | 274.2 | 9.5 | 986 | 6 | CF826586 | CF826586 EST703968 |
| 12 | 273 | 9.5 | 784 | 9 | AZ935182 | AZ935182 BJ_Ba000 |
| 13 | 262.6 | 9.1 | 990 | 7 | CO027797 | CO027797 EST806181 |
| 14 | 261.2 | 9.1 | 684 | 7 | CN907928 | CN907928 030109ABL |
| 15 | 261 | 9.1 | 921 | 6 | CF714848 | CF714848 CCAF343TO |
| 16 | 261 | 9.1 | 946 | 8 | DN809412 | DN809412 73954613 |
| 17 | 259.6 | 9.0 | 713 | 7 | CK757406 | CK757406 atr02-4ms |
| 18 | 258.2 | 9.0 | 689 | 8 | DR440006 | DR440006 EST149_09 |
| 19 | 255.2 | 8.9 | 781 | 8 | DR916924 | DR916924 EST110846 |
| 20 | 254.8 | 8.8 | 863 | 6 | CD375065 | CD375065 PTMM00226 |
| 21 | 254.4 | 8.8 | 945 | 8 | DN808832 | DN808832 76945303 |
| 22 | 249.4 | 8.7 | 757 | 5 | BU027777 | BU027777 QRG7L07.Y |

ALIGNMENTS

RESULT 1
DR637438
LOCUS
DEFINITION DR637438 958 bp mRNA linear EST 11-JUL-2005
EST1028063 FvM Gibberella moniliformis cDNA clone FVMAW64, mRNA
sequence.
ACCESSION DR637438 GI:70712272
VERSION DR637438
KEYWORDS
SOURCE Gibberella moniliformis
ORGANISM Gibberella moniliformis
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
REFERENCE 1 (bases 1 to 958)
AUTHORS Brown,D.W., Cheung,F., Proctor,R.H., Butchko,A.E., Zheng,L, Lee,Y.,
Unterback,T., Smith,S., Feldblyum,T., Glenn,A.E., Plattner,R.D.,
Kendra,D.P., Town,C.D. and Whitelaw,C.A.
TITLE Analysis of 87,000 expressed sequence tags reveals alternatively
spliced introns in multiple genes of the fumonisin gene cluster
JOURNAL Unpublished (2005)
COMMENT Contact: Brown, D.W.
USDA/ARS/NCAUR
USDA
1815 N. University St, Peoria, IL 61604, USA
Tel: 309 681 6230
Fax: 309 681 6689
Email: brown@ncaur.usda.gov
TIGR sequence name: FVMAW64TH
Seq primer: AAT TAA CCC TCA CTA AAG GG.
Location/Qualifiers
1..958
/organism="Gibberella moniliformis"
/mol_type="mRNA"
/strain="m3125"
/db_xref="taxon:117187"
/clone="FVMAW64"
/tissue_type="mycelia"
/clone_lib="FvM"
/note="Vector: pBluescript II SK(+); XR; Site 1: EcoRI;
Site 2: XhoI; anamorph: Fusarium verticillioides. Library
FvM was prepared from pooled RNA obtained from a 48-hour
and a 72 hour, liquid GYM culture from strain M-3125.
Cultures were vacuum filtered and the mycelial mats were
frozen in liquid nitrogen, ground to a powder, and then
added to TRIzol Reagent (Invitrogen, Carlsbad CA) at
approximately 1 g mycelia per 10 ml TRIzol. The cDNA was

directionally ligated into the pBlueScript II SK(+) XR
vector (cDNA Synthesis Kit; Stratagene)."

ORIGIN

Query Match
Best Local Similarity 59.9%; Pred. No. 3.5e-70; Length 958;
Matches 565; Conservative 0; Mismatches 376; Indels 2; Gaps 1;

QY 1209 GATGCGTAAGCCCTCATATGTCGTGGTGTGGGCACAGAGCAGCTAGCAGGGAAGGT 1268
|||||
Db 2 GACGCGTAACAGAGGCGATGTCGTGCGAAGCTGCACGGATGCTGATCAGGTGAAGGA 61
QY 1269 ATGATTATTACCGCTGGGGGAATCGATTACACACCCACTCTCTTCTCCACAAATTC 1328
|||||
Db 62 AAGATTGTGACCGCGCGCTATCGACAGCATATTCATTTATCTGCGCTCAGCAAGTA 121
QY 1329 CCTACCGCTCTAGCAATGCGGTACCAACATGTTGGAGGGGACACAGTCTCTGTAGAT 1388
|||||
Db 122 CCCGAAGCTCTTGCATCTGCTGTAAACCACCATGCTTGGCGGTGGTACCGGCCCAAGTGT 181
QY 1389 GGCACGAATGCGACTACTATCACTCCGGGCAAAATGGAACTTGCCACGCGATGTTCCGCGCA 1448
|||||
Db 182 GGAACGAGCGCAACTACTTGTACGCTGTGTCTCATTAATCGGTCAAATGTTGCAGGCG 241
QY 1449 GCAGAAGAGTATTCTATGAATGTGGCTTTTGGGCAAAAGGCAATAGCTCTAGCAAAAAA 1508
|||||
Db 242 TGCATGACGTTCTCTATCAATATGTTATTAATCTGCAAGAGTAAATGATAGTTCTCTCTGAG 301
QY 1509 CNACTGTAGACNAGTAGAAGGGGCGGATGTTTAAATGATGATGCAAGTGTGCGG 1568
|||||
Db 302 GGTCTGCGCGATCAGGTCAATGCTGTGTGTGCGCTCAAGCTTCATGAGGACTGGGT 361
QY 1569 ACAACACCAAGTGCAGTCACTCTCTGAGCGTGGCAGATGAATACCATGTGCAAGTT 1628
|||||
Db 362 TGCATCTCTGCTGCTATTGACGCTGTCTCAGTGTCTGTGATGAATTCATTAATCAATGT 421
QY 1629 TGTATCCACACCGATACAGTCAATGAGGAGGATTTATGTAGATGACACCCCTAAATGCAATG 1688
|||||
Db 422 CTATTTCACACTGACACGCTTAACGAGTCTGGCTTTGTGCAATCTACGATCGTCTTTC 481
QY 1689 AACGGGCGGCATCCATGCTTACCATCATGAGGAGCGGTGGAGGACACTCACTCTGAT 1748
|||||
Db 482 AAGAAGCCGACAAATTCATATTATCAACAGAGGGTGCAGGAGGTGGCCATGCTCCGGAT 541
QY 1749 GTTATCACCATTGCGAGCGAGTCAATATTTACCTCTCTCCACACCCCTCACTATTCCC 1808
|||||
Db 542 ATCACTCCGTGTAGAGCATCAAAATGTTCTGCCATCATCGACCAACCTCAAGACCA 601
QY 1809 TATACCAATTAATACGGTTGCGAGAACACTTAGACATGCTCATGACATGCCACCACTAGAC 1868
|||||
Db 602 TTCACACGCAATACTCTCGATGAGCATCTCGATATGCTTATGCTGTGCCATCACTTGTCC 661
QY 1869 AAACGATCCGCGAGATTTACAATTTTCTCAAAGCGTATCCGCCCGCGCTCTATCGCG 1928
|||||
Db 662 AAGAATATCCAGAGGATGTAGCTTTCGGCGAGAGCGGTATTGCTGTGAAACCATTTGT 721
QY 1929 GCTGAGATGTGCTCCATGATATGGGTGTGATCGCATGACAGCTCGGATTCGCAAGCA 1988
|||||
Db 722 GCTGAGGATGTATTAACGACAAAGCGCTATCAGCATGATGAGTCTGACTCGCAGGCT 781
QY 1989 ATGGGCGGTGCGAGCGCAAGTGTATCTCGAACTTTGCGAGACTCGGATGAAGATAAAAAA 2048
|||||
Db 782 ATGGGCGGTGCGAGAGGTGCTTTTAAAGAACATGAATACTCGGCATATAAATAGGTG 841
QY 2049 GAATTTGGTAACTTCTCTGAAGATG--GCAAAGATAACGATAATTTCCGATTAAGCGCT 2106
|||||
Db 842 CAGAGGGTGTGTCGGGAGGATGAAGGCACAGGGGCTGATTAATGCGCGTGTAAACGCT 901
QY 2107 ACATCTCCAAATACATATCAACCCCGCTTTTGACCCACGCGT 2149
|||||
Db 902 ATGTCAGCAAGTACTATTAAACCCAGCTATTGCTCAGGCGTT 944

RESULT 2
DR637093
LOCUS
DEFINITION
EST1027718 FvM Gibberella moniliformis cDNA clone FvMAT95, mRNA
sequence.
DR637093
ACCESSION
DR637093.1 GI:70711927
VERSION
KEYWORDS
SOURCE
ORGANISM
Gibberella moniliformis
Gibberella moniliformis
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
REFERENCE
AUTHORS
Brown, D.W., Cheung, F., Proctor, R.H., Butchko, A.E., Zheng, L., Lee, Y.,
Utterback, T., Smith, S., Feldblyum, T., Glenn, A.E., Plattner, R.D.,
Kendra, D.F., Town, C.D. and Whitelaw, C.A.
TITLE
Analysis of 87,000 expressed sequence tags reveals alternatively
spliced introns in multiple genes of the fumonisin gene cluster
JOURNAL
COMMENT
Unpublished (2005)
Contact: Brown, D.W.
USDA/ARS/NCAUR
USDA
1815 N. University St, Peoria, IL 61604, USA
Tel: 309 681 6230
Fax: 309 681 6689
Email: brownw@ncaur.usda.gov
TIGR sequence name: FvMAT95TH
Seq primer: AAT TAA CCC TCA CTA AAG GG.
FEATURES
source
1..946
/organism="Gibberella moniliformis"
/mol_type="mRNA"
/strain="m3125"
/db_xref="taxon:117187"
/clone="FvMAT95"
/tissue_type="mycelia"
/clone_lib="FvM"
/notes="Vector: pBlueScript II SK(+) XR; Site 1: EcoRI;
Site 2: XhoI; anamorph: Fusarium verticillioides. Library
FvM was prepared from pooled RNA obtained from a 48-hour
culture and a 72 hour, liquid GYM culture from strain M-3125.
Cultures were vacuum filtered and the mycelial mats were
frozen in liquid nitrogen, ground to a powder, and then
added to TRIzol Reagent (Invitrogen, Carlsbad, CA) at
approximately 1 g mycelia per 10 ml TRIzol. The cDNA was
directionally ligated into the pBlueScript II SK(+) XR
vector (cDNA Synthesis Kit; Stratagene)."

ORIGIN

Query Match
Best Local Similarity 59.7%; Pred. No. 2.2e-69;
Matches 563; Conservative 0; Mismatches 378; Indels 2; Gaps 1;

QY 1209 GATGCGTAAGCCCTCATATGTCGTGGTGTGGGCACAGAGCAGCTAGCAGGGAAGGT 1268
|||||
Db 2 GACGCGTAACAGAGGCGATGTCGTGCGAAGCTGCACGGATGCTGATCAGGTGAAGGA 61
QY 1269 ATGATTATTACCGTGGGGAATCGATTACACACCCACTCTCTTCTCCACAAATTC 1328
|||||
Db 62 AAGATTGTGACCGCGCGCTATCGACAGCATATTCATTTATCTGCGCTCAGCAAGTA 121
QY 1329 CCTACCGCTCTAGCAATGCGGTGTACCAACCATGTTTGGAGGGGACACAGTCTCTGTAGAT 1388
|||||
Db 122 CCCGAAGCTCTTGCATCTGCTGTAAACCACCATGCTTGGCGGTGGTACCGGCCCAAGTGT 181
QY 1389 GGCACGAATGCGACTACTATCACTCCGGGCAAAATGGAACTTGCCACGCGATGTTCCGCGCA 1448
|||||
Db 182 GGAACGAGCGCAACTACTTGTACGCTGTGTCTCATTAATCGGTCAAATGTTGCAGGCG 241
QY 1449 GCAGAAGAGTATTCTATGAATGTGGCTTTTGGGCAAAAGGCAATAGCTCTAGCAAAAAA 1508
|||||
Db 242 TGCATGACGTTCTCTATCAATATGTTATTAATCTGCAAGAGTAAATGATAGTTCTCTCTGAG 301

```
Qy 1509 CAACTGTAGAACAGTAGAGCGGCGGCGATTGGTTTTAAATTCATGATGAAGACTGGGC 1568
Db 302 GGTCTCGGCGCATCAGGTCAATGCTGTGTGTGGCTCAAGCTTCATGAGGACTGGGGT 361
Qy 1569 ACAACACCAAGTGCATCGATCACTGCTTGAGCGGTGCGACATGAATACGATGTGCAAGTT 1628
Db 362 TGCACATCTCGTCTATGACGGTGTCTCAGTGTCTGTGATGAATTCGATATTCATATGT 421
Qy 1629 TGTATCCACACCCGATCAGTCAATGAGGCGAGGTTATGTAGATGACACCCCTAAATGCAATG 1688
Db 422 CTTATTCACACTGACACGGCTTACGAGTCTGGCTTTGTGCAATCTACGATCGTGTCTTC 481
Qy 1689 AAGCGGCGGCCCATCCATGCTTACCAATGAGGAGCGGTTGAGAGACATCACTACCTGAT 1748
Db 482 AAGAACCGCACAAATTCATCTTATCACAGAGGGTGCAGAGGTGCCCATGTCTCCGAT 541
Qy 1749 GTTATCACCATGGCAGCGAGCTCATATTTCTACCTCTCCACCCACCCCACTATTCCC 1808
Db 542 ATCATCTCGTGGTAGAGCATCAAAATGTTCTGCCATCATCGACCAACCCCTACAAGACCA 601
Qy 1809 TATACCATTAATACGGTTGCGAGAACCTTAGACATGCTCATGACATGCCACCACTAGAC 1868
Db 602 TTCACAGGCAATCTCTCGATGAGCATCTCGATATGCTTATGCTGCTGCCATCACTGTGCC 661
Qy 1869 AAACGCATCCGCGAGGATTTACAATTTCTCAAAGCGGTATCCGCCCGGCTCTATCGCG 1928
Db 662 AAGAATATCCAGAGGATGTAGCTTTCGCGAGAGCGGTATTCGTCTGAAACCATTTGCT 721
Qy 1929 GCTGAAGATGTGCTCCATGATATGGGTGATCGGATGACAGCTCGGATTCGCAAGCA 1988
Db 722 GCTGAGGATGTATTACACGCAAGCGCTATCAGCATGATGAGCTCTGACTCGCAGGCT 781
Qy 1989 ATGGGCGCTGCGAGCGAAGTATTCCTCGAACTTGCAGACTCGGATAGATAAATAA 2048
Db 782 ATGGGCGCTTGGCGAGGTCGTTTAAAGAACATGGAATACTGCGCAATAAATAAGTG 841
Qy 2049 GAATTTGGTAAGCTTCTCGTAAGAT--GGCAAGATAACGATAATTTCCGCAATTAAGCGCT 2106
Db 842 CAGAGGGGTTGGTTGCGGAGGATGAGGCACAGGGGCTCATATGCGCGTGTAAAGCT 901
Qy 2107 ACATCTCCAAATACATATCAACCCCGCTTTGACCCACGCGCT 2149
Db 902 ATGTCAGCAAGTATACTATTATCCCGAGCTATTGCTCAGGGCTT 944

RESULT 3
CNSOABR8 1509 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSLTPGH4ZD08 of Hormone Treated Callus of strain col-0 of
Arabidopsis thaliana (thale cress).
ACCESSION BX816461
VERSION BX816461.1 GI:42471655
KEYWORDS HTC; GSLT cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE
AUTHORS Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
TITLE Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1509)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
```

```
COMMENT The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.

FEATURES
Location/Qualifiers
source 1..1509
/mol_type="mRNA"
/db_xref=taxon:3702"
/clone="GSLTPGH4ZD08"
/tissue_type="Hormone Treated Callus"
/ecotype="Col-0"
/plasmid="pCMVSPORT_6"
gene 1..1509
/gene="Atlg67550"

ORIGIN
Query Match 10.8%; Score 310.2; DB 4; Length 1509;
Best Local Similarity 59.1%; Pred. No. 7.7e-66;
Matches 602; Conservative 0; Mismatches 413; Indels 4; Gaps 4;

Qy 1587 GATCAGCTGCTTGAGCGTGCGAGATGAATACGATGTGCAAGTTTGTATCCACACCGATACA 1646
Db 1 GACAATTTGTTTGGCAGTTTGCAGAGAAATATGATATCCAGGTGAACATTCATATCTGACACC 60
Qy 1647 GTCAATGAGGCGAGTTATGTAGATGACACCTTAATGCAATGCAACGGGCGGCCATCCAT 1706
Db 61 TTAATGAATCCGGCTTTTGTGAGCACACTATCAATGCAATTCGTGGGAGAACATATCAT 120
Qy 1707 GCCTACACATTTGAGGAGCGGGTGAGGACACTCACCTGTGTTATCAACCTGGCAGGC 1766
Db 121 ACATATCACATAAGGTCGTGGTGGAGATGACCAGATATCATTAGAGTTTGTGGA 180
Qy 1767 GAGCTCAATATTTACCCCTCTCCACCAACCCCACTATTTCCCTATACCAATTAATACGGTT 1826
Db 181 GTGAAAATGTACTCCCGTCTATCAACCAACCCCACTATATACGAAAATACTGTA 240
Qy 1827 GCAGAACACTTTAGACATGCTCATGACATGCGCACCACTAGACAAAGCAATCCCGAGGAT 1886
Db 241 GATGAACATCTTGACATGCTGATGTTTGGCCATCACCTTGAACAAGAACATCCCGAAGAT 300
Qy 1887 TTACAATTTTCTCAAGCCGTATCGCCCGGCTCTATCGCGCTGA-AGATGTGCTCCA 1945
Db 301 GTAGCTTTTGTGTAATCAAGAAATAGAGCTGAAACAATAGCTGCGAGAGAGATATATGCA 360
Qy 1946 TGATATGGTGTGATCGCGATGACAAAGCTCGGATTCGCAAGCAATGGGCGGTGCGAGCGCA 2005
Db 361 TGATATGGAGCAATCAGCATTTCTCTCCGACTCAAGCAATGGGTGCGATTGGAGA 420
Qy 2006 AGTGATTTCTCGAACTTGGCAGACTGCGGATAAGAAATAAAGAAATTTGGTAAGCTTCC 2065
Db 421 GGTAATAAGCAGAACTTGGCAGACAGCTGACAAGATGAAGGCTCAGAGGGGAGCCATTAA 480
Qy 2066 TGAAGATGGCAAGATAACGATAATTTCCGCAATTHAGC- GCTACATCTCAATACACTA 2124
Db 481 TCCTAAACATGCTGATGATGATAAATCTCAAGGATCAAGCAGATACATTCGAAAGTACACTA 540
Qy 2125 TCACCCCGCTTTGACCCACCGCTGAGCGAGTATATCGGCTCTCTGTGGAAGGGGCAAGA 2184
Db 541 TAAACCCAGCTATAGCCCAATGGGTTTGAGACTTTAATTTGGCTCGGTGAGGTGAAAGAGC 600
Qy 2185 TCGCCGACTTGGTGGTGGTAATCTGCTCTTTT-TT-GGCGTAAAAACCCAAAAATCGTGATC 2243
Db 601 TGGCTGATCTTGTATATATGGCAGCGCGCTTTCTTCGAGGAGCAAAACCCAGAAATGATAATC 660
```

The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M. URV INRA : Clepet C., Caboche M. Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap. http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full length http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.

```

QY 2244 AAAGCGGTATGGTCTTCTCTGAAATGGGGGATCTTAAACGGTCTGTGCCCACTCCC 2303
D 2244 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 661 AAAGAGGCAATATAGCATGGCTATATGGGCGATGCAAAATGCGAGCATACCCACTCCT 720
QY 2304 CAACCGGTTTATACCGCGAATGTTTGGGCATCAGCGCAAGCGGAAATTTGACACCAGC 2363
D 2304 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 721 GAGCGGTGATATCAAGACCTATGTTTGGAGCAATTTGGAAAGCGCGAAGTGAATACTCT 780
QY 2364 ATCACTTTTGTTCACAAAGTCGCTATGAAATGGCGTGAAGAAAGCTGGGCTTAGAG 2423
D 2364 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 781 GTTGCAATTTGTCAGCAAGCTGCTGTGAGAAAGGGGTAAAGAACTATATGCACTCAAG 840
QY 2424 CCGCAAGTTCTACCGGTCAAAAGTCCGTAACATCACCAAGAAAGACTTCAAGTTCAAC 2483
D 2424 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 841 AAGAGGGTGTAGCGGTGTCCAAACGTGAGCGAGCTCACAAAACCTCGACATGAAGCTGAAT 900
QY 2484 GACAAAACGGCAAAATCACCGTGCATCCGAAACCTTCGAGGTCTTTGTAGATGCAAA 2543
D 2484 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 901 GATGCGTTCCAGAGATCACCGTGGATCCAGAGACATACGTTGTCAAGCAAAATGGCGAG 960
QY 2544 CT-CTGCACTCTAAACCCACCTCGCAAGTGCCTCTAGCCAGCGCTACACTTCTCTCT 2601
D 2544 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 961 GTCTTAGCATGTGCGCCAGCGATTCACTCCCTCTCTGCGGCGGAACTATTTCTCTCT 1019

```

```

RESULT 4
LOCUS CO122599 872 bp mRNA linear EST 16-JUN-2004
DEFINITION GR_Eb04F01.f GR_Eb Gossypium raimondii cDNA clone GR_Eb04F01 5',
mRNA sequence.
ACCESSION CO122599
VERSION CO122599.1 GI:48821286
KEYWORDS EST.
SOURCE Gossypium raimondii
ORGANISM Gossypium raimondii

```

```

REFERENCE 1 (bases 1 to 872)
AUTHORS Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,
Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and
Wing,R.A.
TITLE Global assembly of Cotton ESTs
JOURNAL Unpublished (2004)
COMMENT Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: rwing@genome.arizona.edu
Plate: 04 row: F column: 01.
Location/Qualifiers
1..872
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Eb04F01"
/tissue_type="floral"
/dev_stage="3 to +3 DPA"
/lab_host="DH10B"
/clone_lib="GR_Eb"

```

```

FEATURES
source
1..872
/note="Vector: pCMV.SPORT-6.1; Site_1: NotI; Site_2:
EcoRV; Library made by invitrogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Colonies
plated/picked by AGI. More glycerol clones held in -80."
ORIGIN
Query Match 10.7%; Score 309; DB 7; Length 872;
Best Local Similarity 59.7%; Pred. No. 1.3e-65;
Matches 519; Conservative 0; Mismatches 350; Indels 0; Gaps 0;

```

```

QY 1260 GGGGAAGGTATGATTATTACCGCTGGGGGAATCGATTCAACACCCACTTCTTCTCCA 1319
D 1260 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 2 GGAAGAGGATGATTCGTAACACTGCAGGGGCTATTGACTGTGTGCAATTTTATATGTCTCT 61
QY 1320 CAACAAATTCCTACCGCTCTAGCCAATGGCGTTTACAACCATGTTTGGAGGGCGGACAGGT 1379
D 1320 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 62 CAGTTGGTTCAAGAGCTATTTCAAGTGGCATCAACATTAGTTGGAGGTGGGACCGGA 121
QY 1380 CTTGTAGATGGCAGATCGGACTACTTCACTCCGGGCAAAATGGAACTTTCACCGCATG 1439
D 1380 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 122 CCAGCTGAAGGAACACGTCGCAACTTCTGTAGCGCAGCTCCGTCGCAAAATGAAATGATG 181
QY 1440 TTGGCGCAGCAGAAAGATGATTTCTATGAATGTGGGCTTTTGGGCAAAAGGCAATAGCTCT 1499
D 1440 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 182 CTGCAGTCCACTGATGACTTCCCTCTAAATTTTGGCTTCAAGGGAAGGAAATGGTTCT 241
QY 1500 AGCAAAAAACCACTGTGAGAACAGTAGAAGCGGGCGGATTTGTTTAAATTCGATGAA 1559
D 1500 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 242 AAACCTGAAGAAATACATGAAATAATCAAAGCCGGAGCAATGGGACTGAAACTGCATGAG 301
QY 1560 GACTGGGGCAACACCAAGTGGGATCGATCACTGCTTGGCGTGGCGAGATGAAATACGAT 1619
D 1560 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 302 GATTGGGAACTACACCTGCTGCAATAGACAGTTGTTGGCTGTTGCGAGAACTATATGAT 361
QY 1620 GTGCAAGTTTGTATCCACACCGATACAGTCAATGAGGAGGTTATGTATGATGACACCTA 1679
D 1620 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 362 ATCCAAGTTAATATTCTACTGACACCTTGAAACCAATCTGGATTTGTGGAACAAAT 421
QY 1680 AATGCAATGAACGGCGCGCATCCATGCTTACCACTTGGAGGAGCGGTGGAGACAC 1739
D 1680 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 422 AATGCAATTTAAAGGAAGAACTATTTCACACTTATCACAGTGAAGGTGCGGTTGGTGCAT 481
QY 1740 TCACCTGATGTTATCACCATGGGAGCGAGCTCAATATTCTACCTCTCCACCAACCCCC 1799
D 1740 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 482 GCTCCAGATATCATCAAGATATGTGGGCTTAAAAACGTCCTCCCTTCGTCACAAACCG 541
QY 1800 ACTATTCCCTTATACCAATTAATACGGTTGCGAACAACCTTAGACATGCTCATGACATGCCAC 1859
D 1800 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 542 ACTCGCCCTTATACTTCCAACTACTATAGATGAACATCTTGACATGCTGATGGTTGGCAT 601
QY 1860 CACTAGACAAACGCATCCCGGAGGATTTCAATTTTCTCAAGCCGTATCCGCCCGGC 1919
D 1860 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 602 CACCTCAGCAAGGATATTCTCTGAAGATGTAGCATTTGCAGAAATCAAGGATTAGGGCGGAA 661
QY 1920 TCTATCGCGCTGAAGATGTCTCATGATGGTGTGTCGCGATGACAAAGCTCGGAT 1979
D 1920 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 662 ACCATTGCTCCGGAAGACATATTGCAATGACACAGGGCAATCAGCATTTATTTCTCAGAT 721
QY 1980 TCGCAAGCAATGGGGCGTGCAGGCAAGTGAATTCCTCGAACTTGGCAGACTSCGGATAG 2039
D 1980 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 722 TCACAGCGATGGTTCGCTTCGATTTGGCGAGTGATATGCAGAACTTGGCAAAACGCCCAAG 781
QY 2040 AATAAAAAAGAAATTTGGTAAGCTTCTTGAAGATGGCAAAAGTAACGATAATTTCCGCAATT 2099
D 2040 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 782 ATGAAATACAAAAGGGGACCACTTGGTCTCGAGACTCAGACACGACAACTAAGAAATC 841
QY 2100 AAGCGCTACATCTCCAATACATCAATCA 2128
D 2100 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 842 AAACGTTTATAGCAAAATACACCATAAA 870

```

```

RESULT 5
LOCUS DR633422 925 bp mRNA linear EST 11-JUL-2005
DEFINITION EST1024047 FvM Gibberella moniliformis cDNA clone FVMA534, mRNA
sequence.
ACCESSION DR633422
VERSION DR633422.1 GI:70708256
KEYWORDS EST.
SOURCE Gibberella moniliformis
ORGANISM Gibberella moniliformis
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.

```

REFERENCE 1 (bases 1 to 925)
BROWN, D.W., CHEUNG, F., PROCTOR, R.H., BUTCHKO, A.E., ZHENG, L., LEE, Y.,
UTTERBACK, T., SMITH, S., FELDBLYUM, T., GLENN, A.E., PLATTNER, R.D.,
KENDRA, D.F., TOWN, C.D. and WHITELAW, C.A.
Analysis of 87,000 expressed sequence tags reveals alternatively
spliced introns in multiple genes of the fumonisin gene cluster
Unpublished (2005)
Contact: Brown, D.W.
USDA/ARS/NCAR

1815 N. University St, Peoria, IL 61604, USA
Tel: 309 681 6230
Fax: 309 681 6689
Email: brown@ncar.usda.gov
TIGR sequence name: FVMA534TH
Seq primer: AAT TAA CCC TCA CTA AAG GG.
Location/Qualifiers
1. 925
/organism="Gibberella moniliformis"
/mol_type="mRNA"
/strain="m3125"
/db_xref="taxon:117187"
/clone="FVMA534"
/tissue_type="mycelia"
/clone_lib="FvM"
/notes="Vector: pBlueScript II SK(+); Site 1: EcoRI;
Site 2: XhoI; anamorph: Fusarium verticillioides. Library
FvM was prepared from pooled RNA obtained from a 48-hour
and a 72 hour, liquid YEM culture from strain M-3125.
Cultures were vacuum filtered and the mycelial mats were
frozen in liquid nitrogen, ground to a powder, and then
added to TRIzol Reagent (Invitrogen, Carlsbad, CA) at
approximately 1 g mycelia per 10 ml TRIzol. The cDNA was
directionally ligated into the pBlueScript II SK(+) XR
vector (cDNA Synthesis Kit; Stratagene)."

ORIGIN
Query Match 10.6%; Score 306.2; DB 8; Length 925;
Best Local Similarity 59.9%; Pred. No. 6.7e-65;
Matches 548; Conservative 0; Mismatches 363; Indels 4; Gaps 2;
QY 1312 TTCTCCCAACAATTCCTACCGCTCTAGCCATGCGTTACAAACCATGTTGGAGCG 1371
DB 10 TCTGCCCTCAGCAAGTACCGAAGCTCTTGCACTGGTGTAAACCAACATGTTGGCGTG 69
QY 1372 GCACAGGTCTGTAGATGGACGAATGCGACTACTATCATCTCCGGGCAATGGAATTCG 1431
DB 70 GTACCGGCCCAAGTGTGGAACGAACGCAACTACTTGTACGCTGCTCATTCATCATGC 129
QY 1432 ACCGATGTTGCCGCGCAGCAGAGAGTATTCATGATGTGGCTTTTGGGCAAGGCA 1491
DB 130 GTCAATGTTGCGGGCGTGCATGCACTTCTTATCAATATGTTACTTACGCAAGGTA 189
QY 1492 ATAGCTCTAGCAAAACAACTGTTAGAACAGTAGAGCGGCGGATGTTTAAAT 1551
DB 190 ATGATAGTCTCTGAGGTCTCGCGATCAGTCAATGCTGTGCTGTGGCTCAAGC 249
QY 1552 TGCATGAAGATGGGGCACAACCAAGTGCATCGATCATCTGCTTGAGCGTGGCAGATG 1611
DB 250 TTCATGAGGACTGGGGTGCACCTCCTGCTCTATTGACGCTTGTCTCAGTGTCTGATG 309
QY 1612 AATACATGTGCAAGTTTGTATCCACCGATCAGTCAATGAGGCGAGGTATGATGATG 1671
DB 310 AATTCGATATTCATGCTTATTCACACTGACACGCTTACGAGTCTGGCTTTGTCGAAT 369
QY 1672 ACACCTTAATGAATGAACGGGCGCCATCATGCTACCACTTACGAGGCGGCTG 1731
DB 370 CTACGATGCTGCTTTCAAGAACCGCAAAATTCATCTTATCACAGAGGGTGCAGAG 429
QY 1732 GAGGACATCACCCTGATGTTATCACCATGGCAGCGAGCTCAATATTCACCTCTCTCCA 1791
DB 430 GTGGCCATGCTCGGATATCATCTCCGTGTAGAGCATCAAAATGTTCTGCCATCATCGA 489

1792 CCACCCCACTATTCCCTATACCATTAATACGGTTGAGAACACTTAGACATGCTCATGA 1851
DB 490 CCAACCTTACAGACCACTTTCACAGCAATCTCTCGATGAGCATCTCGATATGCTTATGG 549
QY 1852 CATGCCACCACTAGACAAACGCAATCCGAGGATTTTCAATTTTCTCAAAGCCGATCC 1911
DB 550 TCTGCCATCATTTGTCGAAGATATCCAGAGGATGTAGCCTTCGGGAGAGCCGTTATTC 609
QY 1912 GCCCGCGCTCTATCGCGCTGAAGATGTGCTCCATGATATGGTGTGATCGGATGACAA 1971
DB 610 GTGCTGAACCACTTGTCTGAGGATGTTATACAGCAAAAGCGCTATCAGCATGATGA 669
QY 1972 GCTCGATTCGCAAGCAATGGGCGTGCAGGGAAGTATTCCTCGAATTCGCAACTG 2031
DB 670 GCTCTGACTCGAGGCTATGGCGCTTCGCGAGAGGTGCTTTTAAAGAACATGGAATAC 729
QY 2032 CGGATAGCAATAAAAGAAATTTGTAAGCTTCCTGAAGAT--CGCAAGATAACGATA 2088
DB 730 CGCATAAATAAAGGTGAGAGGGTGTGTTGCCGAGGATGAGGCAACAGGGCTGATA 789
QY 2089 ATTTCCGCAATTAAGCGCTACATCTCCAATACACTATCAACCCCGCTTTGACCCAGCG 2148
DB 790 ATGCGCGTGTAAAGCTATGTCAAGATATATCTATTAAACCCAGCTATTCGACGGCT 849
QY 2149 TGACGAGTATATCGGCTCTCTGGAAGAGGCAAGATCGCGA-CTTGGTGTGTGGAAT 2207
DB 850 TTGGACATGTCTATTTGGAGCATTCAGGTTGGAAGTTTGTCTGATCTTCTTTGGGAT 909
QY 2208 CTGCGCTTTTGGC 2222
DB 910 CCTCGTGGTTGGC 924

RESULT 6
CX676965 748 bp mRNA linear EST 19-JAN-2005
LOCUS ydd29h04.y1 Sea urchin EST Lib1 Strongylocentrotus purpuratus cDNA
DEFINITION clone ydd29h04 5' similar to TR:Q9SR52 Q9SR52 PUTATIVE UREASE. ;
mRNA sequence.
CX676965 GI:57936920
VERSION EST.
KEYWORDS Strongylocentrotus purpuratus
SOURCE Strongylocentrotus purpuratus
ORGANISM Strongylocentrotus purpuratus
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
Echinoidea; Euechinoidea; Echinacea; Echinoida;
Strongylocentrotidae; Strongylocentrotus.
REFERENCE 1 (bases 1 to 748)
AUTHORS Coffman, J.A., Robertson, A.J., Clifton, S., Pape, D., Hillier, L.,
Martin, J., Wylie, T., Dante, M., Meyer, R., Theising, B., Bowers, Y.,
Gibbons, M., Ronko, I., Tsagareishvili, R., Ritter, E., Kennedy, S. and
Wilson, R.
TITLE WashU Sea Urchin EST Project
JOURNAL Unpublished (2004)
COMMENT Contact: Dr. James A. Coffman
WashU Sea Urchin EST Project
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
DNA sequencing by: Washington University Genome Sequencing Center
Seq primer: -28RPPOT
High quality sequence stop: 680.
Location/Qualifiers
1. 748
/organism="Strongylocentrotus purpuratus"
/mol_type="mRNA"
/db_xref="taxon:7668"
/clone="ydd29h04"
/lab host="DH10B"
/clone_lib="Sea urchin EST Lib1"
/note="Vector: pCMVSPORT6.1 (Invitrogen); Site_1: NotI;

Site_2: Smal; Arrayed normalized library of full-length cDNAs representing blastula stage transcriptome of the sea urchin Strongylocentrotus purpuratus, cloned into the vector pCMVSPORT6.1 (Invitrogen)"

| ORIGIN | Query Match | Score | DB 8; | Length | 748; |
|--------|-----------------------|---|--------------------|--------|-----------------------------------|
| | Best Local Similarity | 63.1%; | Pred. No. 6.3e-84; | | |
| | Matches | 466; | Conservative | 0; | Mismatches 273; Indels 0; Gaps 0; |
| QY | 1648 | TCAATGAGCAGGTTATGTAGATGACACCTTAATCAATGAACGCGCGGCCATCATG | 1707 | | |
| Db | | | | | |
| QY | 9 | TCAATGAATCTGCTGTGTGGAGATACCATAGCTCTTTGAGGCGGCACCATCCACA | 68 | | |
| Db | | | | | |
| QY | 1708 | CCTACACATTCAGGAGCGGTGGAGGACACTCACCTGATGTTATCACCATGGCAGGCG | 1767 | | |
| Db | | | | | |
| QY | 69 | CCTATCACTCGAGGCTGCAGGAGGAGCCACGCCCTGACATCATGAAGTAGTGGCG | 128 | | |
| Db | | | | | |
| QY | 1768 | AGCTCAATATTCACCTCTCTGACACCGCCCACTATTCCTATACCATTAATACGGTTG | 1827 | | |
| Db | | | | | |
| QY | 129 | TGCCCAATGTGTACCTCTGTCTACCAACCCACCGCCCTTCACAGTGAACACCATCG | 188 | | |
| Db | | | | | |
| QY | 1828 | CAGAACCTTGACATGCTCATGACATGCCACCTAGACAAACGATCCGGAGATT | 1887 | | |
| Db | | | | | |
| QY | 189 | ATGAACACCTGGATGCTCATGCTGCTGCCACCATCTTGACAAGAAATCTAAAGGAAGCG | 248 | | |
| Db | | | | | |
| QY | 1888 | TACAAATTTCTCAAGCCGTATCCGCCCGGCTCTATCGGGCTGAAGATGTCTCCATG | 1947 | | |
| Db | | | | | |
| QY | 249 | TTGCTTTGACAGTCTCGCATCAGAGCCGAACCATGCTGCTGAAGATATCTTCATG | 308 | | |
| Db | | | | | |
| QY | 1948 | ATATGGGTGTATCGGATGACAAAGCTCGATTCGGAAGCAATGGGGCGTGCAGGCGAAG | 2007 | | |
| Db | | | | | |
| QY | 309 | ATCTGGGAGCAATCAGTATTGTGGCTCGGATTCGCAGGCCATGGGGCGTGTGCGTGAGG | 368 | | |
| Db | | | | | |
| QY | 2008 | TGATCTCTCGACTTGGCAGCTGCGGATGAAGTAAAGAAATTTGGTAAGCTTCCTG | 2067 | | |
| Db | | | | | |
| QY | 369 | TGATCACCCTGATCGGCAGCGGCAGATAAGATGAAGATTTTTCAGAGGCGAGCTGTGAG | 428 | | |
| Db | | | | | |
| QY | 2068 | AAAGTGGCAAGATTAACGATTAATTTCCGCATTAAGCGCTACATCTCCAAATACACTATCA | 2127 | | |
| Db | | | | | |
| QY | 429 | AGGAACGGGTGACACGACACCTTAAGGTCNAGGTTACATGTCGAAGTACACGGTCA | 488 | | |
| Db | | | | | |
| QY | 2128 | ACCCGCTTTGACCCACCGCGTGAGCGAGTATATCGGCTCTGTGGAGAGGCGCAAGATCG | 2187 | | |
| Db | | | | | |
| QY | 489 | ACCCGCGCATTCGCCACGCGCATGGGCCATCTCATTTGGCTCTGTACAGTTGGTAAATGG | 548 | | |
| Db | | | | | |
| QY | 2188 | CCGACTTGTGTGTGGAATCTGCTCTTTTGGCGTAAACCCAAATGCTGATCAAAG | 2247 | | |
| Db | | | | | |
| QY | 549 | CTGACCTGGTTCTTGGAATCCGGCATTTCTTCGAGCTAAACCTGACCTCATCAAAG | 608 | | |
| Db | | | | | |
| QY | 2248 | GGGATATGTTCTTCTCTGAAATGGCGATTTCTAACGGCTGTGCGCCACTCCCAAC | 2307 | | |
| Db | | | | | |
| QY | 609 | GAGTTACATGCTTGGGCACAAATGGGCATGCGCAATGATCATCCCGACCCCTGAAC | 668 | | |
| Db | | | | | |
| QY | 2308 | CGGTTATTACCGCGAAATGTTTGGGCATCACCGGCAAGGCGAAATTTGACACCAAGATCA | 2367 | | |
| Db | | | | | |
| QY | 669 | CTGTCAAGACACAGGAAATGTTTGGTCTTACGGGAAGTCCATCGGTGAGAACTCTGTCA | 728 | | |
| Db | | | | | |
| QY | 2368 | CTTTTGTTCCTCAAGTCG | 2386 | | |
| Db | | | | | |
| QY | 729 | TCTTTGTGTCCTCAAGCAGC | 747 | | |
| Db | | | | | |

| | |
|------------|---|
| RESULT 7 | |
| DR637386 | |
| LOCUS | DR637386 |
| DEFINITION | EST1028011 FvM Gibberella moniliformis cDNA clone FVMAV95, mRNA |
| ACCESSION | DR637386 |
| VERSION | DR637386.1 |
| KEYWORDS | EST. |
| SOURCE | Gibberella moniliformis |
| ORGANISM | Gibberella moniliformis |

Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
1 (bases 1 to 889)
Brown, D.W., Cheung, F., Proctor, R.H., Butchko, A.E., Zheng, L., Lee, Y., Utterback, T., Smith, S., Feldblum, T., Glenn, A.E., Plattner, R.D., Kendra, D.F., Town, C.D., and Whitelaw, C.A.
Analysis of 87,000 expressed sequence tags reveals alternatively spliced introns in multiple genes of the fumonisin gene cluster
Unpublished (2005)
Contact: Brown, D.W.
USDA/ARS/NCAUR
1815 N. University St, Peoria, IL 61604, USA
Tel: 309 681 6230
Fax: 309 681 6689
Email: brownw@ncaur.usda.gov
TIGR sequence name: FVMAV95TH
Seq primer: AAT TAA CCC TCA CTA AAG GG.

Location/Qualifiers
1. .889
/organism="Gibberella moniliformis"
/mol_type="mRNA"
/strain="m3125"
/db_xref="taxon:117187"
/clone="FVMAV95"
/tissue_type="mycelia"
/clone_lib="FvM"
/notes="Vector: pBlueScript II SK(+); XR; Site 1: EcoRI; Site_2: XhoI; anamorph: Fusarium verticillioides. Library FvM was prepared from pooled RNA obtained from a 48-hour and a 72 hour, liquid GYM culture from strain M-3125. Cultures were vacuum filtered and the mycelial mats were frozen in liquid nitrogen, ground to a powder, and then added to TRIzol Reagent (Invitrogen, Carlsbad CA) at approximately 1 g mycelia per 10 ml TRIzol. The cDNA was directionally ligated into the pBlueScript II SK(+) XR vector (cDNA Synthesis Kit; Stratagene)."

| ORIGIN | Query Match | Score | DB 8; | Length | 889; |
|--------|-----------------------|--|--------------------|--------|-----------------------------------|
| | Best Local Similarity | 59.5%; | Pred. No. 3.7e-63; | | |
| | Matches | 524; | Conservative | 0; | Mismatches 353; Indels 3; Gaps 1; |
| QY | 1312 | TTTCTCCACAACTTCCCTACCGCTCTAGCAATGGCTTACCAACCATGTTTGAGGGG | 1371 | | |
| Db | | | | | |
| QY | 10 | TCGCGCTGACGAGTACCGAAGCTCTTGCACTGTGTAAACCACTGTTGCGGTG | 69 | | |
| Db | | | | | |
| QY | 1372 | GCACAGTCTCTGTAGATGGCAACGAATCGACTACTCATCTCCGGGCAAAATGGAACTTGC | 1431 | | |
| Db | | | | | |
| QY | 70 | GTACCGGCCCAAGTGTGTGGAACGAACCACTACTTGTACGCTTGTCTCATACATGC | 129 | | |
| Db | | | | | |
| QY | 1432 | ACCGCATGTTGCGGCGAGCAGAGAGTATTCATGAATGTGGCTTTTGGGCAAGGCA | 1491 | | |
| Db | | | | | |
| QY | 130 | GTCAAAATGTTGCGAGCGTGCAGCTTCATATCAATATTGGTATTACTGCGCAAGGTA | 189 | | |
| Db | | | | | |
| QY | 1492 | ATAGCTCTAGCAAAACAACTTGTAGAACAGTAGAGCGGCGGATTTGTTTTAAT | 1551 | | |
| Db | | | | | |
| QY | 190 | ATGATAGTTCTCTGAGGGTCTGCGCATCAGGTCATGCTGGTCTTGTGCGCTCAAGC | 249 | | |
| Db | | | | | |
| QY | 1552 | TGCATGAAGACTGGGGCACAACACCAAGTGCAGTCGATCACTGCTTGAGCGTGGCAGATG | 1611 | | |
| Db | | | | | |
| QY | 250 | TTCATGAGGACTGGGTTGCACCTCTGCTGCTATTGACGCTTGCTCAGTGTCTGTGATG | 309 | | |
| Db | | | | | |
| QY | 1612 | AATACGATGTGCAAGTTTGTATCCACACCGGATACAGTCAATGAGCAGGTTATGATG | 1671 | | |
| Db | | | | | |
| QY | 310 | AATTCGATATTCAATGTCTTATTACACTGACCGCTTAACGAGTCTGGCTTTGTGCGAAT | 369 | | |
| Db | | | | | |
| QY | 1672 | ACACCTTAATGCAATGAACGGGCGCCATCCATGCTTACCACTTTCAGGAGCGGTG | 1731 | | |
| Db | | | | | |
| QY | 370 | CTAGCATCGCTGCTTTCAAGAACCGCAATTCATATCTATCACAACAGGAGGTCAGGAG | 429 | | |
| Db | | | | | |
| QY | 1732 | GAGGACACTCACTGATGTTTATCAACATGGGCGGAGCTCAATATTCTACCTCTCTCA | 1791 | | |
| Db | | | | | |

Db 430 GTGGCAGATCTCCGGATATCATCTCCGTGGTAGAGCATCAAAATGTTCTGCCATCATCGA 489

Qy 1792 CCACCCCCACTATTCCTTACCATTAATACGGTTTCAGAAACACTTAGACATCTCATGA 1851

Db 490 CCAACCTTACAGACCATTCACACGCAATACTCTCGATGAGCATCTCGATATGCTTATGG 549

Qy 1852 CATGCCACCACTAGACAAACGCATCCGCGAGAGATTTCACAAATTTCTCAAAAGCGTATCC 1911

Db 550 TCTGCCATCACTTGTCCAAGATATCCAGAGGATGTAGCCTTCGCGGAGAGCGGTATTC 609

Qy 1912 GCCCGGCTCTATCCGGCTGAAGATGTCTCCATCATATGGGTGTGATCGCATGACAA 1971

Db 610 GTGCTGAACCAATCTGCTGAGGATGTATTACACGAAAGCGCTATCAGCATGATGA 669

Qy 1972 GCTCGGATTTCGCAAGCAATGGCGGTGCGAGGCGAGTGTCTCTCAACTTTGGCAGACTG 2031

Db 670 GCTCTGACTCGCAGGCTATGGCGGTTCGCGAGAGGTCTGTTTAAAGAAATGAATACTG 729

Qy 2032 CGGATAAGATAAAAAAGAAATTTGGTAACTTCTCTGAAGT---GGCAAGATGAACGATA 2088

Db 730 CGCATAAATAAGTGCAGAGGGTGTGTCGCGAGGATGAGGSCACAGGGGCTGATA 789

Qy 2089 ATTTCGCAATTAAGCGCTACATCTCAAAATACATATCAACCCGCTTTTGAACCCAGCGG 2148

Db 790 ATGCGCGTGTAAAACGCTATGTGAGCAAGTATCTATTAAACCCAGCTATTGCTCAGGCT 849

Qy 2149 TGAGCGAGTATATCGCTCTGTGGAGAGGGCAAGATCGC 2188

Db 850 TTGGACATGTCATTGGAAGCAATTGAGGTGGAAAGTTTGC 889

RESULT 8

LOCUS DR633939

DEFINITION EST1024564 FvM Gibberella moniliformis cdna clone FvMA846, mRNA

ACCESSION DR633939

VERSION DR633939.1 GI:70708773

KEYWORDS EST.

SOURCE Gibberella moniliformis

ORGANISM Gibberella moniliformis

REFERENCE 1. (bases 1 to 796) Brown, D.W., Cheung, F., Proctor, R.H., Butchko, A.E., Zheng, L., Lee, Y., Utterback, T., Smith, S., Feldblyum, T., Glenn, A.E., Plattner, R.D., Kendra, D.F., Town, C.D. and Whitelaw, C.A.

AUTHORS Analysis of 87,000 expressed sequence tags reveals alternatively spliced introns in multiple genes of the fumonisin gene cluster

TITLE Unpublished (2005)

JOURNAL Contact: Brown, D.W.

COMMENT USDA/ARS/NCAUR

USDA 1815 N. University St, Peoria, IL 61604, USA

Tel: 309 681 6230

Fax: 309 681 6689

Email: brown@ncaur.usda.gov

TIGR sequence name: FvMA846TH

Seq primer: AAT TAA CCC TCA CTA AAG GG.

FEATURES

source

1..796

/organism="Gibberella moniliformis"

/mol_type="mRNA"

/strain="M3125"

/db_xref="taxon:117187"

/clone="FvMA846"

/tissue_type="mycelia"

/clone_lib="FvM"

/note="Vector: pBlueScript II SK(+) XR; Site 1: EcoRI; Site 2: XhoI; anamorph: Fusarium verticillioides. Library FvM was prepared from pooled RNA obtained from a 48-hour and a 72 hour, liquid GYAM culture from strain M-3125. Cultures were vacuum filtered and the mycelial mats were

frozen in liquid nitrogen, ground to a powder, and then added to TRIzol Reagent (Invitrogen, Carlsbad CA) at approximately 1 g mycelia per 10 ml TRIzol. The cDNA was directionally ligated into the pBlueScript II SK(+) XR vector (cDNA Synthesis Kit; Stratagene)."

ORIGIN

Query Match 10.3%; Score 296.4; DB 8; Length 796;

Best Local Similarity 60.8%; Pred. No. 1.8e-62;

Matches 483; Conservative 0; Mismatches 311; Indels 0; Gaps 0;

Qy 1243 GCACAGACACTAGCAGCGGAAGGTATGATTATACGGCTGGGGGAATCGATTACACACA 1302

Db 3 GCACGGATGTCGTAGCAGGTGAAGAAAGATTGACCGGGGGCGCTATCCACACGCATA 62

Qy 1303 CCACATTTCTTCTCCACAACAATTCCTCCTACCGCTCTAGCCAAATGGCGTTTACAACCATGT 1362

Db 63 TTCACTTTATTCGCCCTCAGCAAGTACCGAAGCTTTGCAATCTGGTGTAAACCATGTC 122

Qy 1363 TTGGAGCGGCACACAGTCTCTGTAGATGGCAAGTACGACTTACTATCATCTCCGGCAAT 1422

Db 123 TTGGCGGTGTTACCGGCCCAAGTGTCTGAAAGCAACCACTTCTGTACGCTTGGTGCTC 182

Qy 1423 GGAATCTTGACCGCATGTTGGCGGCACAGAGAGTATTCTATGAATGGGCTTTTGG 1482

Db 183 ATTACATGCGTCAAAATGTTGCAGCGCTGCGATCAGTTCCTTATCAATATTGTTACTG 242

Qy 1483 GCAAGGCAATAGTCTCTAGCAAAAAACAATGTTGAGAACAGTAGAGCGGCGCGATTG 1542

Db 243 GCAAGGTAAATGATAGTTCCTCGAGGTCTGCGCGATCAGTCAATGCTGGTCTGTG 302

Qy 1543 GTTTTAAATTCATGAAGACTTGGGGCACAAACCAAGTGGCATCGATCACCTGCTTGAGCG 1602

Db 303 GCCTCAAGCTTCATGAGGACTTGGGTTGCACTCTCTGCTATTGACGCTTGTCTCAGTG 362

Qy 1603 TGCACATGAATACGATGTGCAAGTTTGTATCCACCGCATACAGTCAATGAGGCAAGTT 1662

Db 363 TCTGTGATGAATTCGATATTCAATGTCTTATTCACACTGACACGCTTAAACGAGTCTGCT 422

Qy 1663 ATGTAGATGACACCTAAATGCAATGAACGGGCGGCCATCCATCCCTACCACATTGAGG 1722

Db 423 TTGTCGAATCTACGATCGCTGCTTTCAAGAACCCGCAAAATTCATACTTATCACACAGG 482

Qy 1723 GACGGGTGAGGACACTCACCTGATGTATTCACATGGCAGGCGAGCTCAATATTCTAC 1782

Db 483 GTGCAGAGGTGGCATGCTCCGGATATCATCTCCGTGTGAGAGCATCAAAATGTTCTGC 542

Qy 1783 CTCTCTCCACACCCCACTATTCTCCCTATACCAATTAATACGGTTTGCAGAACACTTAGACA 1842

Db 543 CATCATCGACCAACCCCTACAAGACCAATTCACACGAATACTCTCGATGAGCATCTCGATA 602

Qy 1843 TGCTCATGACATGCCACCACTAGACAAACGATCCGGAGGATTACAAATTTCTCAA 1902

Db 603 TGCTTATGCTGCGCATCACTTGTCCAAGAAATATCCAGAGAGGTGAGCCTTCGCGGAGA 662

Qy 1903 GCGTATCCGCGCGCTCTATCGCGCTGAAGATGTCTCCATGATATGCTGTGATCG 1962

Db 663 GCGTATTCGTGTGAACCAATGCTGCTGAGGATGATTATACACGCAAGGCGGTATCA 722

Qy 1963 CGATGACAAGCTCGGATTTCGCAAGCAATGGGCGCTGCAGGCGAAAGTGAATTCCTCGAACTT 2022

Db 723 GCATGATGACTCTGACTCGCAGGCTATGGCGCTTGGCGAGAGGTGCTGTTTAAAGACAT 782

Qy 2023 GCAGACTCGGAT 2036

Db 783 GGAATACTGCGCAT 796

RESULT 9

LOCUS DR635599

DEFINITION EST1026224 FvM Gibberella moniliformis cdna clone FvMAJ96, mRNA

sequence.

```

ACCESSION DR635599
VERSION 1 GI:70710433
KEYWORDS EST.
SOURCE Gibberella moniliformis
ORGANISM Gibberella moniliformis
REFERENCE Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
AUTHORS Hypocrosmycetidae; Hypocreales; Nectriaceae; Gibberella.
1 (bases 1 to 849)
Brown, D.W., Cheung, F., Proctor, R.H., Butchko, A.E., Zheng, L., Lee, Y.,
Utterback, T., Smith, S., Feldblyum, T., Glenn, A.E., Plattner, R.D.,
Kendra, D.F., Town, C.D., and Whitelaw, C.A.
TITLE Analysis of 87,000 expressed sequence tags reveals alternatively
spliced introns in multiple genes of the fumonisin gene cluster
JOURNAL Unpublished (2005)
COMMENT Contact: Brown, D.W.
USDA/ARS/NCAR
USDA
1815 N. University St, Peoria, IL 61604, USA
Tel: 309 681 6230
Fax: 309 681 6689
Email: brown@ncar.usda.gov
TIGR sequence name: FVMAJ98TH
Seq primer: AAT TAA CCC TCA AAG GG.
FEATURES
source
1..849
/organism="Gibberella moniliformis"
/mol_type="mRNA"
/strain="m3125"
/db_xref="taxon:117187"
/clone="FVMAJ96"
/tissue_type="mycelia"
/clone_lib="FvM"
/notes="Vector: pBlueScript II SK(+). Site 1: EcoRI;
Site 2: XhoI; anamorph: Fusarium verticillioides. Library
FvM was prepared from pooled RNA obtained from a 48-hour
and a 72 hour, liquid GYAM culture from strain M-3125.
Cultures were vacuum filtered and the mycelial mats were
frozen in liquid nitrogen, ground to a powder, and then
added to Trizol Reagent (Invitrogen, Carlsbad, CA) at
approximately 1 g mycelia per 10 ml Trizol. The cDNA was
directionally ligated into the pBlueScript II SK(+) XR
vector (cDNA Synthesis Kit; Stratagene)."
```

Query Match 10.1%; Score 292; DB 8; Length 849;
Best Local Similarity 60.2%; Pred. No. 2.3e-61;
Matches 503; Conservative 0; Mismatches 330; Indels 3; Gaps 1;

```

QY 1317 CCACAACTTCCCTACCGCTCTAGCAATGGCGTTACAAACCATGTTTGGAGCGGCACA 1376
DB 3 CTTGCAAGTACCCGAAGCTTTGGATCTGTCTAACCAACCATGTTGGCGGTGTACC 62
QY 1377 GGTCTGTAGTGGCAAGTGGCACTACTATCACTCCGGGCAAAATGGAACTTGCACCGC 1436
DB 63 GCGCCAAAGTGTGGAAAGCAACCACTACTTGTACGCTGTGCTCATTACATGGCTCAA 122
QY 1437 ATGTTGGCGGAGCAGAAAGATTTCTTAATGTGGCTTTTGGGCAAAAGCAATAGC 1496
DB 123 ATGTTGGAGCGGTGGCATCAGCTTCTATCAATATTGGTATTACTGGCAAGGTAATGAT 182
QY 1497 TCTAGCAAAAACAACTTGTAGAACAGTAGAGCGGCGCATTTGTTTAAATTCAT 1556
DB 183 AGTTCTCTGAGGCTTGGCGGATCAGGTCATCTGTTGCTGTGGCTCAAGCTTCAT 242
QY 1557 GAAGACTGGGGCACAACACCAAGTGGATCGATCGATCTGTTGAGCGTGGCAGATGAATAC 1616
DB 243 GAGGACTGGGTTGCACTCTGCTGCTATTGACGCTTGCTCAGTGTCTGTGATGAATTC 302
QY 1617 GATGTGCAAGTTGTATTCACACCGGATCAGTCAATAGGAGCGGTTATGTAGATGACAC 1676
DB 303 GATATTCAATGTCTATTACACTGACAGCTTAAAGAGTCTGGCTTTGTGCAATCTACG 362
QY 1677 CTAAATGCAATGAACGGCGCGCCATCCATCCATACCACTTGGAGGCGGTGGAGGA 1736
```

```

363 ATCGCTGCTTTCAAGAACCGCACAAATTCATCTATCACACAGAGGGTGCAGGAGTGGC 422
1737 CACTCACCTGATCTTATCACCATGGCGAGGAGCTCAATATCTTACCTCTCCACCAACC 1796
423 CATGCTCCGGATATCATCTCCGGTAGAGCATCAAAATGTTCTGCCATATCAGCAAC 482
1797 CCACATATTCCTTATACCATTAATACGTTTGCAGAACACATTTAGACATGCTCATGACATGC 1856
483 CTTACAAGACCATTCACACGCATATCTCTCGATGAGCATCTCGATATGCTTATGGTCTGC 542
1857 CACCACTAGACAAACGATCCCGAGGATTTCAATTTCTTCAAAGCGTATCGCCCC 1916
543 CATCACTTTGCCAAGAAATATCCAGAGGATGTAGCTTCCGCGAGAGCGGTATTCGTGCT 602
1917 GGCTCTATCGCGGTGAAGATGTCTCCATGATATGGTGTGTGATCGCATGACAGCTCG 1976
603 GAAACCATTGTCTGTGAGGATGATTACACGACAAAGCGCTATCAGCATATGAGCTCT 662
1977 GATTGCAAGCAATGGGGCGTCAGGCGAAGTGAATCTCGAACTTGGCAGACTGCGGAT 2036
663 GACTCGAGGCTATGGCGCTTGGCGAGAGTCTGTTTAAAGAACATGGAATACTGCGCAT 722
2037 AGAATAAAAAAGAAATTTGGTAAAGTTCCTGAAGAT---GGCAAAGATAACGATAATTTTC 2093
723 AAAATAAGGTGCAGAGGGGTGTTGTCGCGAGGATGAGGCGACAGGGGCTGATAATGCG 782
2094 CGCATTAAGCGCTACATCTCCAATACACTATCAACCCCGCTTTGACCCACGGCGT 2149
783 CGTGTAAACGCTATGTCAGCAAGTATACTATTAAACCCAGCTATTGTTCAGGGGCTT 838
```

RESULT 10
CV195867
LOCUS CV195867 766 bp mRNA linear EST 15-SEP-2004
DEFINITION CGF1003441 All Seed coat from mid-season walnut embryos collected
Aug 1 Juglans regia cDNA clone WSC0002_IVF_A11 5', mRNA sequence.
ACCESSION CV195867
VERSION CV195867.1 GI:52124704
KEYWORDS EST.
SOURCE Juglans regia (English walnut)
ORGANISM Juglans regia
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicotyledons;
Rosids; eurosids I; Fagales; Juglandaceae; Juglans.
1 (bases 1 to 766)
Muir, R., Baek, J., Leslie, A., Cook, D. and Dandekar, A.
TITLE Analysis of genes expressed in walnut seed coat tissue
JOURNAL Unpublished (2004)
COMMENT Contact: Abhaya Dandekar, PhD
CAES Genome Facility
UC Davis, Department of Pomology
One Shields Ave, Davis, CA 95616, USA
Tel: 530 752 7784
Fax: 530 752 8502
Email: andandekar@ucdavis.edu
Seq primer: WSCP-TCCGAGATCTGGACGAGC.
FEATURES
source
1..766
/organism="Juglans regia"
/mol_type="mRNA"
/cultivar="Tulare"
/db_xref="taxon:51240"
/clone="WSC0002_IVF_A11"
/sex="Hermaphrodite"
/dev_stage="Mid season fruit collected Aug 1"
/lab_host="XL10-Gold"
/clone_lib="Seed coat from mid-season walnut embryos collected Aug 1"
/notes="Organ: Seed coat; Vector: pTriplEx2; Site 1: SfiIA;
Site 2: SfiIB; Walnut nut samples were harvested from
Tulare trees growing in the 'Stuke Block' in the Wolfskill
experimental orchard located in Winters, California (USA)."

Samples were harvested on August 1, 2001 between 8 and 10 am. Samples were then dissected the same day. Seed coat (pellicle) tissue was separated from embryos and frozen immediately in liquid nitrogen and stored at -80C. A gram of sample was removed and ground to a fine powder in liquid nitrogen. Total RNA was extracted using the hot borate procedure. Poly A+ RNA was obtained using the Poly(A) Purist kit (Ambion). The cDNA library was constructed using the SMART cDNA library kit (Clontech). Primary library was en masse evicted and plasmid DNA containing the cDNA library was isolated from the resultant bacterial population. Plasmid DNA was then transformed into ultra competent E coli cells (XL10 Gold; Stratagene). Transformants were plated out on Q-trays (2000 cfu/tray), picked using a Qbot and archived in 384 well dishes."

ORIGIN

Query Match 9.7%; Score 279.6; DB 7; Length 766;
Best Local Similarity 60.3%; Pred. No. 2.8e-58;
Matches 462; Conservative 0; Mismatches 304; Indels 0; Gaps 0;

QY 1368 GCGGCACAGTCTGTAGATGCGAGCAATGCGACTACTATCACTCCGGGCAAAATGGAC 1427
Db 1 GGGGGAACGGGACCTGCTGATGGAACGCGTGGCACAACCTTGACTCCAGCAATACCAA 60

QY 1428 TTGACCGCATGTTGGCGCACGAGAGTATTCTATGAATGGGCTTTTGGCCAA 1487
Db 61 ATGAAGTTAATGCTGCAATCAACTGATGAGATGCGCTCTAAATTTTGGTTTACGGGAAA 120

QY 1498 GCGAATAGCTCTAGCAAAAAACAACCTTGTAGAACAGTAGAGCGGGCGGATTTGTTT 1547
Db 121 GGGACAGTGCACAAACCTGATGACTACATGAATAATTAGAGCCGGGCAATGGACTG 180

QY 1548 AAATTGCATGAAGACTGGGGCAACACCAAGTGGCATCGATCACTGCTTGAGCGTGGCA 1607
Db 181 AAGCTGATGAGGACTGGGGAACCTACTCTGCTGCAATAGACAATGCTTTGAATTTGCA 240

QY 1608 GATGAATAGATGTGCAAGTTTGTATCCACCGGATACAGTCAATGAGGAGGTTATGTA 1667
Db 241 GAAACAATAGATCCAGGTTAATATCCACCGGACACCTTGAATGAATCGGGATTTGT 300

QY 1668 GATGACACCTAAATGAATGAACGGCGCGCATCCATCGCTPACCACATTTGAGGGAGCG 1727
Db 301 GAACATTTCTATGCTGCAATTAAGAGAACTATTTATACCTTACCACAGTGAAGTGCT 360

QY 1728 GGTGGAGGACACTACCTGATGTTATCACTATGGCAGGCGAGCTCAATATTTCTACCCCTCC 1787
Db 361 GGTGGTGTCACTCCAGATATCATTAAGATATGTGGTGAACAATGTCCTGCCATCA 420

QY 1788 TCCACACCCCACTATTTCCTATACCATTAATAACGGTTGCGAGAACACTTTAGACATGCTC 1847
Db 421 TCTAGGAACCCACACGGCCTTTTACTTTCAACACTATAGATGAGCATCTTGACATGCTG 480

QY 1848 ATGACATGCCACCTAGACAAACGCATCCGGAGGATTTACAAATTTCTCAAACCGGT 1907
Db 481 ATGGTGTGCCATCAGCTCGATAAAGATATTCAGAAAGACGTGCTTTTGTGTAATCAAG 540

QY 1908 ATCCGCGCGGCTCTATCCGGCTGAAGATGTGCTCCATGATGGGTGTGATCGCGATG 1967
Db 541 ATAAGGGTGAACAAATGCTCGAGAGGATATTTTCATGATATGAGGGCAATTAGCATC 600

QY 1968 ACAAGCTCGAATTCGCAAGCAATGGGGCGTGCAGGCGAAAGTATTCCTCGAATTCGGCAG 2027
Db 601 ATTGCTTCGATGCAACAGCTATGGGTGCGATTGGAGAGGTGATTAGCAGAACTTGGCAA 660

QY 2028 ACTGCGGATAAGATAAAAAAGAAATTTGTAAGCTTCCTGAAGATGGCAAGATAACGAT 2087
Db 661 ACTGCCCAAGATGAAGTTTACAAAGAGGGTCAATTTGACCTTAGTGGGACCAACAATGAC 720

QY 2088 AATTTCCGATTAAGCGCTACATCTCCAATATACACTATCAACCCCG 2133
Db 721 AATCTTGTATCAAGCGTTACATTTGCTAAATACACAATAATCTCTG 766

RESULT 11
CF826586/c

LOCUS
DEFINITION

CF826586 986 bp mRNA linear EST 01-APR-2004
EST703988 Coccidioides posadasii saprobic phase cDNA library, 2 to 4 kb Coccidioides posadasii cDNA clone CDB361 3' end, mRNA sequence.

ACCESSION
VERSION

CF826586.1 GI:45932643

KEYWORDS
SOURCE

Coccidioides posadasii
Coccidioides posadasii
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Onygenales; mitosporic Onygenales; Coccidioides.

REFERENCE
AUTHORS

Gardner,M.J. and Cole,G.T.

TITLE

Analysis of gene expression in Coccidioides posadasii mycelia and spherules via expressed sequence tags

JOURNAL
COMMENT

Unpublished (2003)
Other ESTs: EST703969

CONTACT: Gardner MJ

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301 838 3519

Fax: 301 838 0208

Email: gardner@igr.org.

FEATURES
source

Location/Qualifiers

1..986

/organism="Coccidioides posadasii"

/mol_type="mRNA"

/strain="C735"

/db_xref="taxon:199306"

/clone="CDB361"

/dev_stage="saprobic phase (mycelia)"

/lab_host="E. coli DH10B, T1 phage resistant"

/clone_lib="Coccidioides posadasii saprobic phase cDNA library, 2 to 4 kb"

/note="Vector: pExpress 1; Site_1: Not 1; Site_2: Eco RV;

Coccidioides posadasii saprobic phase cDNA library, size fractionated cDNA 2 to 4 kb"

ORIGIN

Query Match 9.5%; Score 274.2; DB 6; Length 986;
Best Local Similarity 59.1%; Pred. No. 6.6e-57;
Matches 512; Conservative 0; Mismatches 343; Indels 12; Gaps 2;

QY 1721 GGGAGCGGTGGAGGACACTCACCTGATGTTATCACCATGGCAGCGAGCTCAATATTCT 1780
Db 986 GGGTGTGGAGAGGCCACGCTCCAGATATCATATCCGTCGCGAGAGCAACGTCCT 927

QY 1781 ACCCTCTCCACACCCCACTATTCCCTATATACCATTAATACGGTTGAGAACACTTAGA 1840
Db 926 GCCCAGCAGTACGAATCCCACTCGTCGTATACGGTAAATACTTTAGATGAACATCTGA 867

QY 1841 CATGCTCATGACATGCCACCACTAGACAAACGATCCGCGAGGATTTACAAATTTCTCA 1900
Db 866 CATGGTAATGCTGCGCATCATTTGTCCAAAGATATTCTCTGAAGACGTGGCTTTTTCGGA 807

QY 1901 AAGCCGTATCCGCGCCCGCTCTATCCGCGGTGAAGATGTGCTCCATGATATGGGTGTGAT 1960
Db 806 AAGCCGATCCGATCCGAGACATTTGCTGCAGAGAGCGTTCTTCATGACCGGAGCCAT 747

QY 1961 CGCGATGACAAGCTCGGATTCGCAAGCAATGGGCGGTGCGAGCGAGGTGATTCCTCGAAC 2020
Db 746 CAGCATGCTATCTCCGACTCTCAAGCTATGGGACGCTGTGGAGAGTTGTTGTTGCGGAC 687

QY 2021 TTGGCAGACTCGGATGAAGATAAAGAAATTTGGTAAGCTTCTCTGAAGT--GGCAA 2077
Db 686 ATGGAACTCTGCATGAAGATAAAGAAATTTGGAAACGAGGGCGACTCAAGGAAGATGAAGGGAC 627

QY 2078 AGATAACGATAATTTCCGCGATTAGCGCTACATCTCCAAATACACTATCAACCCCGCTTT 2137
Db 2137

The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
Fax: 301 838 0208
Email: gardner@tigr.org

FEATURES
source

1. .990
Location/Qualifiers
/organism="Coccidioides posadasii"
/mol_type="mRNA"
/strain="C735"
/db_xref="taxon:199306"
/clone="CIFAM57"
/dev_stage="spherules"
/lab_host="E. coli DH10B, T1 phage resistant"
/clone_lib="Coccidioides posadasii spherule cdna library,
0.5 to 5.3 kb"
/notes="Vector: pExpress 1; Site 1: Not 1; Site 2: Eco RV;
Coccidioides posadasii spherule cdna library, 0.5 to 5.3
kb"

ORIGIN

Query Match 9.1%; Score 262.6; DB 7; Length 990;
Best Local Similarity 58.8%; Pred. No. 5.2e-54;
Matches 525; Conservative 0; Mismatches 339; Indels 29; Gaps 3;

QY 1696 GGGCCATCGCTACCACTTGGAGGCGGTGGAGGACACTCACCTG-ATGTTATC 1754
DB 990 GAGTCACTTACAGTACCACTAGAGGTGCTGGAGAGGCCACGTCAGATATCAT 931

QY 1755 ACCATGGCAGGAGCTCAATATTCACCTCTCCACACCCCCACTATTCCTATACC 1814
DB 930 TCCGTCGTCGAGAGCAAGCTCTGCCACGACGATCGAATCCCACTGTCGTATACG 871

QY 1815 ATTAATACGGTTGCAGAACACTTAGACATGCTCATGACATGCCACACCTAGACAAACGC 1874
DB 870 GTAAATACTTTAGATGAACATCTGGACATGGTAATGCTCTGCCCATCATTTGTCCAAAGAT 811

QY 1875 ATCCGCGAGGATTTACAAATTTCTCAAAGCGTATCGCCCGGCTCTATCGGGCTGAA 1934
DB 810 ATCTCTGAAGAGCTGGCTTTTGGGAAGCCGATCCGATCCGAGCAATTTGCTGCAGAA 751

QY 1935 GATGTGCTCCATGATATGGGTGTGATCGGATGACAAAGCTCGGATTCGCAAGCAATGGGG 1994
DB 750 GAGGTTCTTCATGACACGGGAGCCATCAGCATGCTATCTCGGACTCTCAAGCTATGGGA 691

QY 1995 CQTGACGGCGAAGTATCTCTGAACTTGGCAGACTGCGGATGAAGAAATAAAAAGAAATTT 2054
DB 690 CGCTGTGGAGAAGTTGTTGTCGGACATGGAACTGCAATGAAGAAATAAAAATGGAACGA 631

QY 2055 GGTAAAGCTTCCTGAAGAT---GGCAAGATAAGATAATTCGCAATTAAGCGCTACATC 2111
DB 630 GGGCGACTCAAGGAAGATGAAGGGACGGAATCTGTAATTTTATAGGGTTAAACGGGTATATC 571

QY 2112 TCCAAATACATATCAACCCCGCTTTGACCCAGCGGTGAGCAGTATATCGGCTCTGTG 2171
DB 570 ACCAAGTACACCATCAACCTCGCATTTGCACAGGGATGGCCACACTATTGGAGCGTG 511

QY 2172 GAAGAGGGCAAGATCCCGCACTTGGTGGTGTGGAATCTCTGCTTTTGGCGTAAACCC 2231
DB 510 GAAATGGCAAGACCGCTGATTTGTTCTGTGGAAATTTTGCCAACTTTGGGACTAAACCG 451

QY 2232 AAAATCGTATCAAGCGCGTATGGTGGTCTTCTGTAATGGCGATTTCTAACGGGTCT 2291
DB 450 AGTATGGTCTTGAAGTCTGGAATGGGTGTCTCAGCGCAGATGGGTGATCCCAATGGCTCT 391

QY 2292 GTGCCACTCCCAACCGGTTTATTACCGGAAATGTTGGGACATCACGCGAGGCGAAA 2351
DB 390 ATCCCAACAAATCGAGCTTATTTATGAGGCTTATGTACGCTCTCTCA----- 344

QY 2352 TTTGACACGAGCATCATCTTTGTTTCCAAAGTCGCTATGAAATGGCGTGAAGAAAG 2411
DB 343 -----ATCATGTTCTGATCCCAAGCATCCATCAAGCTTGTGATCATCGACAGT 296

QY 2412 CTGGGCTTAGAGCCGCCAAGTTCTACCGGTCAAAAACCTCCGTAAATACATCACCAAGAAC 2471
DB 295 TACCACCTGAAGAAGCGGATCGAGCCAGTGAAGATTTGTCGGAATATTAAGCAAGAGAT 236

QY 2472 TTCAAGTTTCAACGACAAAACGCAAAATACCGTCCGATCCGAAACCTTCCGAGGTCTTT 2531
DB 235 ATGAATTTATGATATTTATGCCCAAAATGAGATCGATCCGAGAGCTATGTTGTCGAG 176

QY 2532 GTAGATGGCAAACTCTGCACCTCTAAACCCACCTCGCAAGTGCCTCTTAGCCCCA 2584
DB 175 GCTGACGGGAGAGTGCACCGCTGAGCCAGTGTCCAGATTTGCCCTTTAACACA 123

RESULT 14
CN907928
LOCUS 684 bp mRNA linear EST 07-JUN-2004
DEFINITION 030109ABLC001286HT (ABLC) Braeburn cell culture three days after
subculture Malus x domestica cdna clone ABLC001286, mRNA sequence.
CN907928
VERSION
KEYWORDS
SOURCE
ORGANISM
Malus x domestica
Malus x domestica
EST.
CN907928.1 GI:48380429

REFERENCE
AUTHORS
Beuning, L., Bowen, J., Crowhurst, R., Gleave, A., Janssen, B.,
McArtney, S., Newcomb, R., Ross, G., Snowden, K., Walton, E. and Yauk, Y.
HortResearch Apple EST Project
Unpublished (2004)
Contact: Gleave, A.
Sequencing Facility
The Horticulture and Food Research Institute of New Zealand Ltd
120 Mt Albert Rd, Mt Albert, Auckland, New Zealand
Tel: 00 64 09 815 4200
Fax: 00 64 09 815 4201
Email: est@hortresearch.co.nz.

FEATURES
source
1. .684
/organism="Malus x domestica"
/mol_type="mRNA"
/db_xref="taxon:3750"
/clone="ABLC001286"
/tissue_type="Cell cultures"
/dev_stage="Exponential phase of growth (three days after
subculture)"
/clone_lib="(ABLC) Braeburn cell culture three days after
subculture"
/note="Vector: pBluescript SK(-); Library sequenced by
Genesis Research & Development"

ORIGIN

Query Match 9.1%; Score 261.2; DB 7; Length 684;
Best Local Similarity 61.4%; Pred. No. 1.1e-53;
Matches 419; Conservative 0; Mismatches 263; Indels 0; Gaps 0;

QY 1546 TTAATTTGATGAAGACTGGGGCAACACCAAGTGGCATGATCATCTGCTTGAGCGTGG 1605
DB 3 TGAAGCTGCATGAGGACTGGGGAAGCACTCTGCTGCAATCGACAATTTGTTGGCTGTTG 62

QY 1606 CAGATGAATAGGATGTGCAAGTTTGTATCCACACCGATACAGTCAATGAGCGAGTTAG 1665
DB 63 GAGATCAATATGACATCCAGGTAAACATCCATACAGACACCTTGAATGAATCTGGAATTG 122

QY 1666 TAGATGACACCTTAAATGCAATGAACGGCGCGCCATCCATGCCCTTACACATTTGAGGGAG 1725
DB 123 TAGAGCATCAATTTGCTGCAATTTAAGGAAGAACTATCCACTTACCACAGTGAAGTG 182

QY 1726 CGGTTGGAGGACACTCACCTGATGTTATCCACATGGCAGGCGAGCTCAATATTTACCCCT 1785
DB 183 CAGGTGGGGGCCATGCTCCAGATATATCAAAAGTCTCGGGTGTGAAAAATGTCTCTGCCAT 242

QY 1786 CCTCACACCCCGCCTATTCCTATACCATTAATACGGTTGCCAGAACACTTAGACATGC 1845
Db CATCAACGAATCCCAACCGGCTTTTCACCTCAAAATACCTATAGATGAGCATCTTGATATGC 302
QY 1846 TCATGACATGCCACACCTAGACAAACGATCCGCGAGGATTTACAAATTTCTCAAAGCC 1905
Db TGATGGTTTGCATCACCTTGACAAGGACATTCAGAAGATGTAGCTTTTGCTGAATCA 362
QY 1906 GTATCCGCCCGCGCTCTATCCGGCTGAAGATGTCTCCATGATATGGGTGTGATCGCGA 1965
Db GAATAAGGCTGAACAATTCGTCAGAAGATATTTTGCACGATATGGGGCCATTAGCA 422
QY 1966 TGACAACTCGGATTCGGAAGCAATGGGGCGTGCAGGCGAAGTATTCCTCGAATTTGGC 2025
Db TTGTATCTTCTGATTCACAGGCTATGGGTGCGATTTGGAGAGGTGATAATCAAGAACTTGGC 482
QY 2026 AGACTCGCGGTAAGATTAAGAAATTTGGTAACTTCTCCTGAAGTGCAGAAATAGC 2085
Db AAACAGCTGACAAAGATGAATCACAAAGGGTGTGATAGAACTAGTGGATCCGCAATG 542
QY 2086 ATAATTTCCGCAATTAAGCGCTACATCTCCAAATACATCAACCCCGCTTTGACCCACG 2145
Db ACAACTTCGGATCAACGATACATCGCAAGTACACTATAATCCAGCAATAGTTATG 602
QY 2146 GCGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCAAGATCCCGACTTTGGTGTGGA 2205
Db GATTTTCTCAGTATGTGGTCTATTTAGGTGGGAAGCTGCTGATCTGTCTATGGA 662
QY 2206 ATCTCGCTTTTTCGGCGTAAA 2227
Db AGCCATCGTCTTTTGGCGCAAA 684

RESULT 15
LOCUS CF714848/C
DEFINITION CCF343TO C.neoformans strain JEC21 Cryptococcus neoformans var.
ACCESSION CF714848
VERSION CF714848.1 GI:41569007
KEYWORDS EST.
SOURCE Cryptococcus neoformans var. neoformans (Filobasidiella neoformans
var. neoformans)
ORGANISM Cryptococcus neoformans var. neoformans
Eukaryota; Fungi; Basidiomycota; Hymenomycetes;
Heterobasidiomycetes; Tremellomycetidae; Tremellales; Tremellaceae;
Filobasidiella.
REFERENCE 1 (bases 1 to 921)
AUTHORS Loftus,B.
TITLE End sequencing of clones from a Full length enriched, normalized
JEC21 cDNA library
JOURNAL Unpublished (2003)
COMMENT Other_ESTs: CCF3433TR
Contact: Brendan Loftus
The Institute for Genomic Research (TIGR; www.tigr.org)
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-3543
Fax: 301-838-0208
Email: bjoftus@tigr.org
Seq primer: RF.

FEATURES
source
1..921
/organism="Cryptococcus neoformans var. neoformans"
/mol_type="mRNA"
/strain="JEC21"
/db_xref="taxon:40410"
/clone="CCF343"
/clone_lib="C.neoformans strain JEC21"
/note="Vector: pCMVSPORT6; Site_1: NotI_EcoRV; The full
length, normalized library was prepared from a variety of
conditions using RNA provided by Joseph Heitman and
Jennifer Lodge"
ORIGIN

Query Match 9.1%; Score 261; DB 6; Length 921;
Best Local Similarity 55.4%; Pred. No. 1.3e-53;
Matches 504; Conservative 0; Mismatches 405; Indels 0; Gaps 0;
QY 1694 GCGCGCATCCATCCCTACACATTTGAGGGAGCGGTGGAGGACACTCACTGATGTTAT 1753
Db 921 GAGGACTATTCACAGTTACACACTGAAGGAGTGTGGCGGACATGCTCAGACATCAT 862
QY 1754 CACCATTGGCAGGCGAGCTCAATATTTACCTCTCCACACCCCTACTATTCCTATAC 1813
Db 861 CGTTGCTGTGAATATGAATAACGCTTTGCTAGTTCACCAACCCACTAGGCTTATGC 802
QY 1814 CATTAATACGGTTGCGAAGACATTTAGACATGCTCATGACATGCCACACCTTAGACAAACG 1873
Db 801 TGTAAACACTCTTGATGAGCATCTTGACATGCTTATGCTGTGTCACCACTCGATAAGTC 742
QY 1874 CATCCGCGAGGATTTACAATTTTCTCAAGCGGTATCCGCCCGCTCTATTCGCGGCTGA 1933
Db 741 TATCCCGGAGGACATTTGCCCTTTGCCGACTCTCGTATCCGTTCTGAACCCGTTGAGCGGA 682
QY 1934 AGATGTGCTCCATGATATGGGTGTGATCGGATGACAAAGCTCGGATTCGCAAGCAATGGG 1993
Db 681 AGAGGTTTACAGGACACAGGCGGATTTCCATGATCTCATCTGACTGTCAAGCTATGGG 622
QY 1994 GCGTGAGGCGAGTGAATTCCTCGAACTTGGCGAGACTGCGGATGAAGATAAAAAGAAAT 2053
Db 621 TCGTATCGGTGAAGTCACTCATCGTACATGCGGTACGGTTCGGAAGATGAAGCAATTCG 562
QY 2054 TGTGAAGCTTCTGAAGATGGCAAGATAAGATAATTTCCGCATTAAGCGCTACATCTC 2113
Db 561 TGGTCTCTTGGGCGGTGAACCCACGAGAGACAAATAGGGTGAACCGTATGTTGC 502
QY 2114 CAAATACACTATCAACCCCGCTTTTGACCCAGCGGTGAGGATATCGGCTCTGTGGA 2173
Db 501 CAAATACACTATTAACCCCGCTTTTGACCCAGGATATGTCGCACTCATCTGCTCAAGTTGC 442
QY 2174 AGAGGCAAGATCGCGACTTGGTGTGGAATCTCTGCGCTTTTGGCGGTAAACCCAA 2233
Db 441 CGTGGCTGTCTGCGAGATCTGGTCTTCTGACGCGAGAGTCTTTTGGTCTAGACCGGA 382
QY 2234 AATCGTATCAAAAGCGGTATGGTGTCTTCTGAAATGGGCGATTTAAACGGTCTGT 2293
Db 381 GATGATCTCAAGGGTGGTGTCTATTTGTTGGGTGCAATGGGCGATGCGAACGCTTCGAT 322
QY 2294 GCCCACTCCCAACCGGTTTATACCGGAAATGTTTGGGCATCACGCGAGGCGAAAT 2353
Db 321 CCCCACTGTGACGCGCTCATTTGGCAGGCCCATGTGGGGTCTCAGCCTGAGGCGCGTGC 262
QY 2354 TGACACAGCATCACTTTTGTTCMAAAGTCGCTATGAAATGGCGTGAAGAAAGAGCT 2413
Db 261 ACTCAATTCATTTGTTGGGTGAGCGAGGATCTCTTGACAGGATCTCGTGAAGAGATT 202
QY 2414 GGGCTTAGAGCGCAAGTTCTCAACGGTCAAAAACCTGCGTAACATCAACAAAGAGATT 2473
Db 201 CAACATCAAGAAAGGCGGAGGCTGTCAAGAACTGTGTCGAATTTGAAAGAGGATAT 142
QY 2474 CAAGTTCAACGACAAAACGCAAAAATCACCGTCGATCCGAAACCTTCGAGGCTTTGT 2533
Db 141 GAAAGTGAATGACAGTATGCCGAAGTGAATGCTGTCGATCCGAGACCTACGACGTTACGC 82
QY 2534 AGATGCAAACTCTGCACTCTTAACCCACCTCGCAAGTGTCTTAGCCCGAGCGCTACAC 2593
Db 81 TGAGGCGCTTTTGGACGCTCCACCGGAGACAAATCCCACTGACCAAGAGATCTT 22
QY 2594 TTTCTTCTA 2602
Db 21 CGTTTACTA 13

Search completed: November 29, 2005, 07:20:55
Job time : 10217 secs

XX
PS
Claim 1. SEO ID NO 1: 76pp: English.

XX The present invention relates to a novel *Helicobacter felis* urease X and
CC Y subunit polypeptides and immunogenic fragments. The polypeptides are
CC used in the manufacture of vaccines against *Helicobacter felis* infections
CC and in diagnostic tests to detect antibodies against *Helicobacter felis*.
CC *Helicobacter felis* is difficult to grow so it is more convenient to use
CC the expression products of the genes encoding the urease X and Y subunits
CC in the manufacture of vaccines. The present sequence represents an urease
CC subunit polypeptide complex encoding sequence of the invention.

XX Sequence 2883 BP; 871 A; 642 C; 666 G; 698 T; 0 U; 6 Other;

Query Match 99.9%; Score 2880.6; DB 7; Length 2883;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2883; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| QY | 1 | RGGRAGATTTTCARCACTTCAAGCACATATGATCCTGTGTGGTGGTAAATTCR | 60 |
| DB | 1 | RGGRAGATTTTCARCACTTCAAGCACATATGATCCTGTGTGGTGGTAAATTCR | 60 |
| QY | 61 | ACTTGTAAATCTATTATTAATTTTAAATTAATTAATTAATTAATTAATTAATTA | 120 |
| DB | 61 | ACTTGTAAATCTATTATTAATTTTAAATTAATTAATTAATTAATTAATTAATTA | 120 |
| QY | 121 | TTACTTATATAAAAGTTAATAAAAGTAACGAAATTAGGACTATAATCCCATTCCTT | 180 |
| DB | 121 | TTACTTATATAAAAGTTAATAAAAGTAACGAAATTAGGACTATAATCCCATTCCTT | 180 |
| QY | 181 | TAAATTTTAACAAGGAGTAATAGTGAACTACACCCAAAGACCAAGAAAGTTCTT | 240 |
| DB | 181 | TAAATTTTAACAAGGAGTAATAGTGAACTACACCCAAAGACCAAGAAAGTTCTT | 240 |
| QY | 241 | GTATATATGCGGGGAGTGCTAGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCA | 300 |
| DB | 241 | GTATATATGCGGGGAGTGCTAGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCA | 300 |
| QY | 301 | ACCGAAGCCATTGCTTACATTAGTGCCCATATTATGAGCAGAGCGCGGTGGAATAA | 360 |
| DB | 301 | ACCGAAGCCATTGCTTACATTAGTGCCCATATTATGAGCAGAGCGCGGTGGAATAA | 360 |
| QY | 361 | AACCGTTGCCAGCTTATGGAAGAGTGCACTGCACTTTTGAATAAAGATGAATGCC | 420 |
| DB | 361 | AACCGTTGCCAGCTTATGGAAGAGTGCACTGCACTTTTGAATAAAGATGAATGCC | 420 |
| QY | 421 | CGGGTGGGTAAATGTTTCCCATCTAGGTGTAGAGCACCTTTCCTGATGTTACGAA | 480 |
| DB | 421 | CGGGTGGGTAAATGTTTCCCATCTAGGTGTAGAGCACCTTTCCTGATGTTACGAA | 480 |
| QY | 481 | ACTTGTAACTGTGAATTTGGCCCATCGAACCCAGATGAGCACTTCAAGCGGGCGAAGTGA | 540 |
| DB | 481 | ACTTGTAACTGTGAATTTGGCCCATCGAACCCAGATGAGCACTTCAAGCGGGCGAAGTGA | 540 |
| QY | 541 | ATTGGTTCCGATAAAGACATCGAGCTCAATGAGGCAAGAAAGTAACCGAACTTGAGGT | 600 |
| DB | 541 | ATTGGTTCCGATAAAGACATCGAGCTCAATGAGGCAAGAAAGTAACCGAACTTGAGGT | 600 |
| QY | 601 | TACTAATGAAGGCGCTAAATCTTGCATGTGGTAGGCATTTCCACTTTTGAAGCTAA | 660 |
| DB | 601 | TACTAATGAAGGCGCTAAATCTTGCATGTGGTAGGCATTTCCACTTTTGAAGCTAA | 660 |
| QY | 661 | CAAGGCATTAATAATTCGATCGTGAATAAGGCTATGCAACCGCTAGATATTCCTCTGG | 720 |
| DB | 661 | CAAGGCATTAATAATTCGATCGTGAATAAGGCTATGCAACCGCTAGATATTCCTCTGG | 720 |
| QY | 721 | CAACAGCTACGATTTGGGCGAGCAAAACCCGCAAAAGTGCAGTTGATTCCTCTGGTGG | 780 |
| DB | 721 | CAACAGCTACGATTTGGGCGAGCAAAACCCGCAAAAGTGCAGTTGATTCCTCTGGTGG | 780 |
| QY | 781 | CAGTAAAAAGTGAATGGCATGAACGGGCTTGTGAATAACATCGCGATGAACGCCATAA | 840 |
| DB | 781 | CAGTAAAAAGTGAATGGCATGAACGGGCTTGTGAATAACATCGCGATGAACGCCATAA | 840 |
| QY | 841 | ACATAAAGCGCTTGACAAGCGGAAATCTCACGGATTTTCAAGTAAGGAGACTCCCATGA | 900 |

| | | | |
|----|------|---|------|
| DB | 841 | ACATAAAGCGCTTGACAAGCGGAAATCTCAGGATTTTATCAAGTAAGGAGACTCCCATGA | 900 |
| QY | 901 | AAATGAAAAAACAAGAAATATGTAATACTACGACCCCAACAAAGGGGATAAAGTGCGCT | 960 |
| DB | 901 | AAATGAAAAAACAAGAAATATGTAATACTACGACCCCAACAAAGGGGATAAAGTGCGCT | 960 |
| QY | 961 | TAGGAGATACCGATCTTTGGGCGAAGTAGAAATGACTATACCACTTATGCGGAAGAAC | 1020 |
| DB | 961 | TAGGAGATACCGATCTTTGGGCGAAGTAGAAATGACTATACCACTTATGCGGAAGAAC | 1020 |
| QY | 1021 | TTAAATTTGGCGGGGTAAATCTATCCGTGAGGGTATGGGTGAGAGCAATAGCCCTGATG | 1080 |
| DB | 1021 | TTAAATTTGGCGGGGTAAATCTATCCGTGAGGGTATGGGTGAGAGCAATAGCCCTGATG | 1080 |
| QY | 1081 | AAAAACCCCTAGATTTAGTTCATCACTAAACCGATGATTTCGACTACACCGGGATTTACA | 1140 |
| DB | 1081 | AAAAACCCCTAGATTTAGTTCATCACTAAACCGATGATTTCGACTACACCGGGATTTACA | 1140 |
| QY | 1141 | AAGCCGACATTTGGGATTTAAAAACGGCAAAATCCATGGCATTTGGCAAGCAGAAACAAGG | 1200 |
| DB | 1141 | AAGCCGACATTTGGGATTTAAAAACGGCAAAATCCATGGCATTTGGCAAGCAGAAACAAGG | 1200 |
| QY | 1201 | ACATGCAAGATGGCGTAAGCCCTCATATGTCGTGGGTGGGCAAGAAAGCACTAGCAG | 1260 |
| DB | 1201 | ACATGCAAGATGGCGTAAGCCCTCATATGTCGTGGGTGGGCAAGAAAGCACTAGCAG | 1260 |
| QY | 1261 | GGGAAGGTATGATTTATACCGCTGGGGGAATCGATTACACACCCACTTCCTTCTCCAC | 1320 |
| DB | 1261 | GGGAAGGTATGATTTATACCGCTGGGGGAATCGATTACACACCCACTTCCTTCTCCAC | 1320 |
| QY | 1321 | AACAAATTCCTACCGCTTAGCCAAATGCGGTTACAAACATGTTTGGAGCGGCAAGGTC | 1380 |
| DB | 1321 | AACAAATTCCTACCGCTTAGCCAAATGCGGTTACAAACATGTTTGGAGCGGCAAGGTC | 1380 |
| QY | 1381 | CTGTAGATGSCAAATGCGACTACTACTCTCGGCAAAATGGAATTCACCGCATGT | 1440 |
| DB | 1381 | CTGTAGATGSCAAATGCGACTACTACTCTCGGCAAAATGGAATTCACCGCATGT | 1440 |
| QY | 1441 | TGCGCGCAGCAGAAAGATATTCTATGATGTTGGGCTTTTGGGCAAGGCAATAGCTCTA | 1500 |
| DB | 1441 | TGCGCGCAGCAGAAAGATATTCTATGATGTTGGGCTTTTGGGCAAGGCAATAGCTCTA | 1500 |
| QY | 1501 | GCAAAAAACAATTTGTAGAACAGTAGAAGCGGCGGATTTGTTTAAATTTGATGAAG | 1560 |
| DB | 1501 | GCAAAAAACAATTTGTAGAACAGTAGAAGCGGCGGATTTGTTTAAATTTGATGAAG | 1560 |
| QY | 1561 | ACTGGGSCACAAACCAAGTGCATCGATCACTGCTTGGCGTGGCAGATGAATACGATG | 1620 |
| DB | 1561 | ACTGGGSCACAAACCAAGTGCATCGATCACTGCTTGGCGTGGCAGATGAATACGATG | 1620 |
| QY | 1621 | TGCAAGTTTGTATCCACACCGATACAGTCAATGAGGAGGTTATGTAGATGACACCTTAA | 1680 |
| DB | 1621 | TGCAAGTTTGTATCCACACCGATACAGTCAATGAGGAGGTTATGTAGATGACACCTTAA | 1680 |
| QY | 1681 | ATCAATGAACGGGCGGCGCATCCATGCTTACCTACCACTTGGAGGCGGTTGGAGGACT | 1740 |
| DB | 1681 | ATCAATGAACGGGCGGCGCATCCATGCTTACCTACCACTTGGAGGCGGTTGGAGGACT | 1740 |
| QY | 1741 | CACCTGATGTTATCACCATGGCAGGCGAGTCAATATTTCTACCTCTCCACCAACCCCA | 1800 |
| DB | 1741 | CACCTGATGTTATCACCATGGCAGGCGAGTCAATATTTCTACCTCTCCACCAACCCCA | 1800 |
| QY | 1801 | CTATTCCTATACATTTAATAACCGTTGAGAACTTAGACATGCTCATGATGCCACC | 1860 |
| DB | 1801 | CTATTCCTATACATTTAATAACCGTTGAGAACTTAGACATGCTCATGATGCCACC | 1860 |
| QY | 1861 | ACCTAGACAAACCGATCGGCGAGATTTAAATTTTCTCAAGCGGTTATCGCCCGGCT | 1920 |
| DB | 1861 | ACCTAGACAAACCGATCGGCGAGATTTAAATTTTCTCAAGCGGTTATCGCCCGGCT | 1920 |
| QY | 1921 | CTATCGGCGTGAAGATGTCCTCATGATATGGGTGTGATCGGATGAACAGCTCGGAT | 1980 |
| DB | 1921 | CTATCGGCGTGAAGATGTCCTCATGATATGGGTGTGATCGGATGAACAGCTCGGAT | 1980 |

Db 1921 CTATCGCGCTGAAGATGTGCTCCATGATATGGGTGTGATCGCGATGCAAGCTCGGATT 1980
Qy CGCAAGCAATGGGCGTGCAGGCGAAGTGATTCCTCGAACTTGGCAGACTCGGATAAGA 2040
Db 1981 CGCAAGCAATGGGCGTGCAGGCGAAGTGATTCCTCGAACTTGGCAGACTCGGATAAGA 2040
Qy 2041 ATAAAAAGAAATTTGGTAAAGCTTCTGGAAGTGCAGAAAGATAACGATAATTTCCGCATTA 2100
Db 2041 ATAAAAAGAAATTTGGTAAAGCTTCTGGAAGTGCAGAAAGATAACGATAATTTCCGCATTA 2100
Qy 2101 AGCGTACATCTCCAAATACATATCAACCCGCTTTGACCCAGCGGTGAGCAATATA 2160
Db 2101 AGCGTACATCTCCAAATACATATCAACCCGCTTTGACCCAGCGGTGAGCAATATA 2160
Qy 2161 TCGCTCTGTGGAAGAGGCGAAGATCGCCGACTTGGTGTGGAATCCTGCCCTTTTGG 2220
Db 2161 TCGCTCTGTGGAAGAGGCGAAGATCGCCGACTTGGTGTGGAATCCTGCCCTTTTGG 2220
Qy 2221 GCSTAAACCCAAATCGTGATCAAGGCGGTATGGTGTCTCTGAAATGGCGGATT 2280
Db 2221 GCSTAAACCCAAATCGTGATCAAGGCGGTATGGTGTCTCTGAAATGGCGGATT 2280
Qy 2281 CTAAACGGCTGTGCCACCTCCCAACCGGTTTATACCGGAAATGTTGGGCATCAG 2340
Db 2281 CTAAACGGCTGTGCCACCTCCCAACCGGTTTATACCGGAAATGTTGGGCATCAG 2340
Qy 2341 GCAAGGCGAATTTGACACAGCATCACTTTGTTTCCAAAGTCGCTATGAAATGGCG 2400
Db 2341 GCAAGGCGAATTTGACACAGCATCACTTTGTTTCCAAAGTCGCTATGAAATGGCG 2400
Qy 2401 TGAAGAAGAGCTGGGCTTAGAGCGCAAGTTCTACCGGTCAAAAACCTGCCGTAAACATCA 2460
Db 2401 TGAAGAAGAGCTGGGCTTAGAGCGCAAGTTCTACCGGTCAAAAACCTGCCGTAAACATCA 2460
Qy 2461 CCAAGAAAGACTTCAAGTTCAAGCAGCAAAACGGCAAAAATCACCGTCGATCCGAAACCT 2520
Db 2461 CCAAGAAAGACTTCAAGTTCAAGCAGCAAAACGGCAAAAATCACCGTCGATCCGAAACCT 2520
Qy 2521 TCAGGCTTTGTAGATGGCAACTCTGCACCTTAACCCACCTCGCAAGTGCCTTAG 2580
Db 2521 TCAGGCTTTGTAGATGGCAACTCTGCACCTTAACCCACCTCGCAAGTGCCTTAG 2580
Qy 2581 CCAGAGCTACATTTCTTAGGCACAATGCCCTTTGGGGCAGGTTATTTAGGAA 2640
Db 2581 CCAGAGCTACATTTCTTAGGCACAATGCCCTTTGGGGCAGGTTATTTAGGAA 2640
Qy 2641 TCTTCATCAACACGCTGCAATCGGTCTTTGCGTGTGCGATCGTGTGCTTTAAACAAC 2700
Db 2641 TCTTCATCAACACGCTGCAATCGGTCTTTGCGTGTGCGATCGTGTGCTTTAAACAAC 2700
Qy 2701 TTTTCATCTTAAGCAATCCCAATTTTAAATTAATTTAAATCTTATTAATTAATTAAT 2760
Db 2701 TTTTCATCTTAAGCAATCCCAATTTTAAATTAATTTAAATCTTATTAATTAATTAAT 2760
Qy 2761 TAGCCCCCTCATTTTAAAGGAGAATATGCGTAGGTCTTTGCTATTTGCTATGCGGG 2820
Db 2761 TAGCCCCCTCATTTTAAAGGAGAATATGCGTAGGTCTTTGCTATTTGCTATGCGGG 2820
Qy 2821 TTGTTGGTGTGCTGGGCGCAAGGGTATTGAAACCCATCGCTTCAAAAAGTAGAAGCCAC 2880
Db 2821 TTGTTGGTGTGCTGGGCGCAAGGGTATTGAAACCCATCGCTTCAAAAAGTAGAAGCCAC 2880
Qy 2881 AGG 2883
Db 2881 AGG 2883

RESULT 2
ADJ58249
ID ADJ58249 standard; DNA; 2452 BP.
XX
AC
ADJ58249;
XX

DT 06-MAY-2004 (first entry)
XX
DE UreaseXY subunit encoding sequence #4.
XX
KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis; ds.
XX
OS Helicobacter felis.
FH Key Location/Qualifiers
CDS 48..728
FT /*tag= a
FT /product= "urease protein"
FT 739..2445
FT /*tag= b
FT /product= "urease protein"
XX
PN EP1176192-A2.
XX
PD 30-JAN-2002.
XX
PF 11-JUL-2001; 2001EP-00202666.
XX
PR 17-JUL-2000; 2000EP-00202565.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Kusters JG, Cattoli G;
XX
PS WPI; 2002-124384/17.
DR P-PSDB; ADJ58250, ADJ58251.
XX
PT Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
the diagnosis of Helicobacter felis infections and in the preparation of
vaccines.
XX
PS Disclosure; SEQ ID NO 13; 76pp; English.
XX
CC The present invention relates to a novel Helicobacter felis urease X and
Y subunit polypeptides and immunogenic fragments. The polypeptides are
used in the manufacture of vaccines against Helicobacter felis infections
and in diagnostic tests to detect antibodies against Helicobacter felis.
CC Helicobacter felis is difficult to grow so it is more convenient to use
the expression products of the genes encoding the urease X and Y subunits
in the manufacture of vaccines. The present sequence represents an
ureaseXY subunit of the invention.
XX
SQ Sequence 2452 BP; 746 A; 568 C; 590 G; 548 T; 0 U; 0 Other;
Query Match 77.3%; Score 2228; DB 7; Length 2452;
Best Local Similarity 94.3%; Pred No. 0;
Matches 2312; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
Qy 159 AGGACTATAATCCCATTCCTTTAAATTTTAAACAGGAGTAATAGGTGAAACTCACAC 218
Db 1 AGGACTATAATCCCATTCCTTTAAATTTTAAACAGGAGTAATAGGTGAAACTCACAC 60
Qy 219 CCAAGAGCAAGAAAAGTTCTTTATATTATTCGGGGGAAGTGGTAGAAGCGCAAG 278
Db 61 CCAAGAGCAAGAAAAGTTCTTTATATTATTCGGGGGAAGTGGTAGAAGCGCAAG 120
Qy 279 CAGAGGCTTAAAGCTCAACCAACCGCATTTACATTAGTGCCTATTTATG 338
Db 121 CAGAGGCTTAAAGCTCAACCAACCGCATTTAGCTTACATTAGTGCCTATTTATG 180
Qy 339 ACGAAGCGCGCTGGGAAAAAACCCTTGTGCCAGCTTATGGAAGAGTGCATGCACTTTT 398
Db 181 ACGAGGCGCTGCTGGCAAAAACCCGTTGCGAACTTATGGAAGAGTGCATGCACTTTT 240
Qy 399 TGAAGAAAAGATGAAGTAATGCGCGGGTGGGTAATATGTTTCCCGATCTAGGTGTAAG 458
Db 241 TGAAGAAAAGAGAGGTGATGTCGCGGGTGGGAATATGTCCTGATTTGGGCGTGAAG 300
Qy 459 CCACCTTCTCTGATGGGTACGAAACTTGTAACTGTGAAATGGCCCATCGAACCAGATGAGC 518

Db 1321 AGAAGCGGCGGATTTGGCTTTAAATTTGATGAAGACTGGGCGACAACACCAAGTCCGAT 1380
QY CGATCACTGCTTTGACGCTGGCAGATGAATACATGTGCAAGTTTGTATCCACACCGATAC 1645
Db 1381 CGATCACTGCTTTGACGCTAGCAGATGAATACATGTGCAAGTTTGTATCCACACCGATAC 1440
QY 1646 AGTCAATGAGGCGGTTATGTAGATGACACCTAAATGCAATGAACGGCGCGCCATCCA 1705
Db 1441 GGTCAATGAGGCGGTTATGTAGATGACACCTAAATGCGATGAACGGCGCGCCATCCA 1500
QY 1706 TGCCTTACCACATTTAGGCGGCGGTGGAGGACATCACTACCTGTATTTATCACCATGCGAG 1765
Db 1501 TGCCTTACCACATTTAGGCGGCGGTGGAGGACATCACTACCTGTATTTATCACCATGCGAG 1560
QY 1766 CGAGCTCAATATTTCTCAACCTCTCCACACCCCGCTCTATTCCTATACCATTAATACGGT 1825
Db 1561 CGAGCTCAATATTTCTCAACCTCTCCACACCCCGCTCTATTCCTATACCATTAATACGGT 1620
QY 1826 TGCAGAACACTTTAGACATGCTCATGACATGCGACACCTAGACAAACGGATCCGCGAGGA 1885
Db 1621 TGCAGAACACTTTAGACATGCTCATGACCTGCGACACCTAGACAAACGGATCCGCGAGGA 1680
QY 1886 TTTACAAATTTCTCAAGCCGTATCGGCCCGGCTCTATCGCGCTGAAGATGTGCTCCA 1945
Db 1681 TCTCCAGTTTCTCCAAAGCCGTATCGGCCCGGCTCTATTCGCGCTGAAGATGTGCTCCA 1740
QY 1946 TGATATGGGTGTGATCGGATGACAGCTCGGATTCGCAAGCAATGGGCGTGCAGCGGA 2005
Db 1741 TGATATGGGTGTGATCGGATGACAGCTCGGATTCGCAAGCAATGGGCGGCTGCAGGA 1800
QY 2006 AGTGATTTCTCGAATCTTGGCAGACTGCGGATGAAGAAATAAAAAAGAAATTTGGTAAGCTTCC 2065
Db 1801 AGTGATTTCTCGAATCTTGGCAGACTGCGGATGAAGAAATAAAAAAGAAATTTGGTAAGCTTCC 1860
QY 2066 TGAAGATGCAAGATGAACGATAATTTCCGATTAAGCGCTACATCTCMAAATACACTAT 2125
Db 1861 TGAAGATGCGAGATGAATGACAACTTCGCGCATCAAAACGCTATATCTCAAATACACCAT 1920
QY 2126 CAACCCCGCTTTGACCCAGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGAAGAT 2185
Db 1921 TAAATCCCGCTTTGACCCATGGCGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGAAGAT 1980
QY 2186 CGCCGACTTGTGTGTGGAATCTGCTCTTTTGGCGTAAACCCCAAAATCTGTGATCAA 2245
Db 1981 CGCCGACTTGTGTGTGGAATCTGCTCTTTTGGTGTAAACCCCAAAATCTGTGATCAA 2040
QY 2246 AGGCGTATGTGTGTCTCTGAAATGGCGGATTTAAACGCTCTGTGCGCCACTCCCA 2305
Db 2041 AGGCGGATATGTGTGTCTCTGAAATGGCGGATTTAAACGCTCTGTGCGCCACTCA 2100
QY 2306 ACGGTTTATACCGGAAATGTTGGGATCAGCGCAAGCGCAAAATTTGACACCCAGCAT 2365
Db 2101 GCGCGTTTATACCGGAAATGTTGGGATCAGCGCAAGCGCAAAATTTGACACCCAGCAT 2160
QY 2366 CACTTTTGTTCCTCAAGTGCCTTATGAAATGGCGTGAAGAAAGCTTGGGCTTTAGAGCG 2425
Db 2161 CACTTTTGTTCCTCAAGTGCCTTATGAAATGGCGTGAAGAAAGCTTGGGCTTTAGAGCG 2220
QY 2426 CCAAGTTCTACCGGTCAAAACCTGCGTAAACATCACCAAGAAAGACTTCAAGTTCAACGA 2485
Db 2221 CAAGGTGCTACCGGTCAAAACCTGCGCAACATCACTAAGAAAGACTTCAAAATTCACAA 2280
QY 2486 CAAACGGCAAAATACACCGTCGATCCGAAACCTTCGAGGCTTTGTAGATGCGCAACT 2545
Db 2281 CAAGCGGCGCATATCACTGTGATCTTAAACCTTCGAGGCTTTGTAGATGCGCAACT 2340
QY 2546 CTGCACTCTAAACCCACCTCGCAAGTCTCTAGCCCGCTACACTTTCTTCTAGGC 2605
Db 2341 CTGCACTCTAAACCCCGCTCTGAAAGTGCTCTAGCCCAACGCTACACTTTCTTCTAGGC 2400
QY 2606 ACAAT 2610
Db 2401 ACAAT 2405

RESULT 4
ADJ58246

ID ADJ58246 standard; DNA; 2407 BP.

XX AC ADJ58246;

XX DT 06-MAY-2004 (first entry)

XX DE UreaseXY subunit encoding sequence #3.

XX KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis; ds.

XX OS Helicobacter felis.

XX FH Key Location/Qualifiers
CDS 2..682FT FT /*tag= a
/product= "urease protein"

FT FT 693..2399

FT FT /*tag= b
/product= "urease protein"

XX PN EP1176192-A2.

XX XX 30-JAN-2002.

XX XX 11-JUL-2001; 2001EP-00202666.

XX XX 17-JUL-2000; 2000EP-00202565.

XX XX (ALKU) AKZO NOBEL NV.

XX XX Kusters JG, Cattoli G;

XX DR WPI; 2002-124384/17.

XX DR P-PSDB; ADJ58247, ADJ58248.

XX PT Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
the diagnosis of Helicobacter felis infections and in the preparation of
vaccines.

XX PS Disclosure; SEQ ID NO 10; 76pp; English.

XX CC The present invention relates to a novel Helicobacter felis urease X and
Y subunit polypeptides and immunogenic fragments. The polypeptides are
used in the manufacture of vaccines against Helicobacter felis infections
and in diagnostic tests to detect antibodies against Helicobacter felis.
XX CC Helicobacter felis is difficult to grow so it is more convenient to use
the expression products of the genes encoding the urease X and Y subunits
in the manufacture of vaccines. The present sequence represents an
ureaseXY subunit of the invention.

XX SQ Sequence 2407 BP; 720 A; 574 C; 584 G; 528 T; 0 U; 1 Other;

Query Match 74.9%; Score 2158.6; DB 7; Length 2407;

Best Local Similarity 93.6%; Fred. No. 0;

Matches 2251; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

QY 206 GTGAAACTCACACCCCAAGAGCAAGAAAGTTCTGTATATATATCGCGGCGAAGTGGCT 265

Db 2 GTGAAACTCACACCCCAAGAGCAAGAAAGTTCTGTATATATATCGCGGCGAAGTGGCT 61

QY 266 AGAAAGCGCAAGAGAGGGCTTAAAGCTCAACCAACCGAAGCCATTGCTTACATTAGT 325

Db 62 AGAAAGCGCAAGAGGGCTTAAAGCTCAACCAACCGAAGCCATTGCTTACATTAGT 121

QY 326 GCCCATATATGAGGAAGCGCGCTGGAAAAAAGAAACCGTTGCCAGCTTATGGAAGAG 385

Db 122 GCCCATATATGAGGAAGCGCGCTGGAAAAAAGAAACCGTTGCCAGCTTATGGAAGAG 181

QY 386 TGCATGCACTTTTGAAGAAAGATGAAGTATGCCGCGGTGGTAAATATGGTTCCCGAT 445

[illegible]

| | | |
|-------|--|-------|
| 13262 | GAATGTGGGCTTTTGGGCAAGCGAATAGCTCCAGTAAAAAAACAACCTTGTAGAACAAAT | 13221 |
| Qy | AGAAAGCGGCGCGCATGGTTTAAAATTGCAAGAAGACTGGGGCAACAACCAAGTCGCAAT | 15851 |
| | | |
| | | |
| 13222 | AGAAAGCGGCGCGCATCGGCTTTAAAATTGCAAGAAGACTGGGGCACAACCTCCAAGTGCAAAT | 13811 |
| Qy | CGATCACTGCTTGAGCGGTGGCAGATGAATACGATGTGCAAGTTTGTATCCAACCGATAC | 15451 |
| | | |
| | | |
| 13821 | CGATCACTGCTTGAGCGGTAGCAGATGAATACGATGTGCAAGTTTGTATCCAACCGATAC | 14411 |
| Qy | AGTCAATGAGGCAAGGTTATGTAGATGACACCCCTAAATGCAATCGAAAGCGGCGGCCCATCCA | 17051 |
| | | |
| | | |
| 1442 | GGTCAATGAGGCAAGGTTATGTAGATGACACCCCTGAAATGCGATGAAAGCGGCGGCCCATCCA | 15011 |
| Qy | TGCTCTACCAATTTGAGGAGCGGCTGGAGACACTCACCTGATGTTTATCACCATGCGCAGG | 17651 |
| | | |
| | | |
| 1502 | TGCTCTACCAATTTGAGGAGCGGCGGAGGACACTCACCTGATGTTTATCACCATGCGCAGG | 15611 |
| Qy | CGAGCTCAATATTTCTACCTCTCCACACCCGCCACTATTTCCCTCTATACCATTTAATACGGT | 18251 |
| | | |
| | | |
| 15621 | CGAGCTCAATATTTCTACCTCTCCACACCCGCCACTATTTCCCTCTATACCATTTAATACGGT | 16211 |
| Qy | TGCAGAACACTTAGACATGCTCATGACCTGCCACCACTAGACAAACGCGATCCGCGAGGA | 18851 |
| | | |
| | | |
| 1622 | TGCAGAACACTTAGACATGCTCATGACCTGCCACCACTAGATAAAACGCGATCCGCGAGGA | 16811 |
| Qy | TTTACAAATTTCTCAAAAGCCGATTCGCGCCGCGCTCTATCGCGGCTGAAGATGTGCTCCA | 19451 |
| | | |
| | | |
| 1682 | TTTACAAATTTCCAAAGCCGATTCGCGCCGCGCTCTATCGCGGCTGAAGATGTGCTCCA | 17411 |
| Qy | TGATATGGGTGTGATTCGCGATGACAAGCTCGGAATTCGCAAGCAATGGGCGGTGCAGGCGA | 20051 |
| | | |
| | | |
| 1742 | TGATATGGGCGTGATTCGCGATGACAAGCTCGGAATTCGCAAGCAATGGGCGGTGCAGGCGA | 18011 |
| Qy | AGTGATTCCTCGAACTTTGGCAGACTCGGATAGAAATAAAAGAAATTTGGTAAAGCTTCC | 20651 |
| | | |
| | | |
| 1802 | AGTGATTCCTCGAACTTTGGCAGACTCGGATAGAAATAAAAGAAATTTGGTAAAGCTTCC | 18611 |
| Qy | TGAAGATGGCAAAAGATAACGATAAATTTTCGCAATTAAGCGCTACATCTCCAAAATACACTAT | 21251 |
| | | |
| | | |
| 1862 | TGAAGATAGTGCAAGATAACGAACTTTCCGTATCAAAACGCTACATCTCCAAATACACTAT | 19211 |
| Qy | CAACCCCGCTTTGACCCACCGCGTGACGAGTATATCGGCTCTGTGGAAGAGGCGCAAGAT | 21851 |
| | | |
| | | |
| 1922 | TAACCCCGCTCTAAACCCATCGGGCTAAGCGAGTATATCGGCTCTGTGGAAGAGGCGCAAAAT | 19811 |
| Qy | CGCGCACTTGCGGTGTGGAAATCCTCGCTTTTGGCGTAAAAACCCAAAATCGTGAATCAA | 22451 |
| | | |
| | | |
| 1982 | CGCTGATTTTGGTGTGGAAATCCTCGCTTTTGGTGTGAAAACCTTAAGATTTGTGATCAA | 20411 |
| Qy | AGCGGTGTAGGTGCTTCTCTGAAATGGCGAATCTTAAGCGCTGTGCGCACTCCCA | 23051 |
| | | |
| | | |
| 2042 | AGCGGTGTAGGTGCTTCTCTGAAATGGCGAATCTTAAGCGCTGTGCGCTTACACTCA | 21011 |
| Qy | ACCGGTTTATTACCGGGAATGTTTGGGCATCACGGCAAGGCGAAATTTTGACACCAAGCAT | 23651 |
| | | |
| | | |
| 2102 | GCCGGTTTATTACCGGGAATGTTTGGGCATCACGGCAAGGCGAAATTTTGACACCAAGCAT | 21611 |
| Qy | CATTTTGTTCAAAGTGCCTATGAAAATGGCGTGAAGAAAGCAAGCTGGGCTTAGAGCG | 24251 |
| | | |
| | | |
| 2162 | CACTTTGTTCCTCAAAAGTCGCTTATGAAAATGGCGTGAAGAAAGCAAGCTTAGAGCG | 22211 |
| Qy | CCAAGTTCCTACCGGTCAAAAACCTCGCGTAACATCAACAAGAAGACCTCAAGTTCAACGA | 24851 |
| | | |
| | | |
| 2222 | CAAGGTGCTACCGGTGAAAACCTGCGCGAATCATCTAAGAAGACCTTCAATTTCAACAA | 22811 |
| Qy | CAAAAACGGCAAAATCACCGTCGATCCGAAAACTTTTCGAGGCTCTTTGTAGATGGCAAACT | 25451 |
| | | |
| | | |
| 2282 | CAAGACGGCGCATATCACTGCTCGATCCTTAAACCTTCGAGGCTCTTTGTAGATGGCAAACT | 23411 |
| Qy | CTGCACTCTTAAACCCCACTTCGCAAGTGCGCTCTAGGCCAGCGGTACACTTTCTTAGGC | 26051 |
| | | |
| | | |
| 2342 | CTGCACTCTTAAACCCCGCTCTGAAGTGCGCTCTAGGCCAGCGGTACACTTTCTTAGGC | 24011 |

```
QY      2606 ACAATG 2611
      |||||
Db      2402 NCAATG 2407

RESULT 5
ADJ58243
ID      ADJ58243 standard; DNA; 2183 BP.
XX
AC      ADJ58243;
XX
DT      06-MAY-2004 (first entry)
XX
DE      UreaseXY subunit encoding sequence #2.
XX
KW      immunogenic fragment; antibacterial; Vaccine; Helicobacter felis; ds.
XX
OS      Helicobacter felis.
XX
FH      Key
FT      CDS
FT      Location/Qualifiers
FT      3..683
FT      /*tag= a
FT      /product= "urease protein"
FT      694..2181
FT      /*tag= b
FT      /product= "urease protein"
XX
PN      EP1176192-A2.
XX
PD      30-JAN-2002.
XX
PF      11-JUL-2001; 2001EP-00202666.
XX
PR      17-JUL-2000; 2000EP-00202565.
XX
PA      (ALKU ) AKZO NOBEL NV.
XX
PI      Kusters JG, Cattoli G;
XX
PI      WPI; 2002-124384/17.
DR      P-PSDB; ADJ58244, ADJ58245.
XX
PT      Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
PT      the diagnosis of Helicobacter felis infections and in the preparation of
PT      vaccines.
XX
PS      Disclosure; SEQ ID NO 7; 76pp; English.
XX
CC      The present invention relates to a novel Helicobacter felis urease X and
CC      Y subunit polypeptides and immunogenic fragments. The polypeptides are
CC      used in the manufacture of vaccines against Helicobacter felis infections
CC      and in diagnostic tests to detect antibodies against Helicobacter felis.
CC      Helicobacter felis is difficult to grow so it is more convenient to use
CC      the expression products of the genes encoding the urease X and Y subunits
CC      in the manufacture of vaccines. The present sequence represents an
CC      ureaseXY subunit of the invention.
XX
SQ      Sequence 2183 BP; 638 A; 505 C; 550 G; 490 T; 0 U; 0 Other;
Query Match      67.6%; Score 1948; DB 7; Length 2183;
Best Local Similarity 93.3%; Pred. No. 0;
Matches 2035; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY      206 GTGAACTCACCCAAAGAGCAAGAAAGTCTTGTATATATATGCGGGCGAAGTGCT 265
      |||||
Db      3 GTGAACTCACCCAAAGAGCAAGAAAGTCTTGTATATATATGCGGGCGAAGTGCT 62
      |||||
QY      266 AGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCGAGCCATTGCTTACATTAGT 325
      |||||
Db      63 AGAAAGCGCAAGCAGAGGGCTTAAAGCTCAATCAACCGAGCCATTGCTTACATTAGT 122
      |||||
QY      326 GCCCATATTATGACGAGCGCGCGTGGAAAAAACCCTTGCCAGCTTATGGAAGAG 385
      |||||
```

```
Db      |||||
123 GCCCATATTATGACGAGCGCGCGTGGAAAAAACCCTTGCTGTAACCTTATGGAAGAA 182
      |||||
QY      386 TGCATGCACCTTTTGAAGAAAGATGAAGTAAATCCCGGGTGGGTAAATATGTTTCCCGAT 445
      |||||
Db      183 TGTATGCACCTTTTGAAGAAAGATGAGTAAATCCCGGTGTTGGGAATATGTTTCCCTGAT 242
      |||||
QY      446 CTAGGTGTAGAAGCACCTTTTCTGTATGGTACGAAACTTGTAACTGTGAATTTGGCCCATC 505
      |||||
Db      243 TTGGCGGTAGAAGCCACTTTCCCGATGGCACCACAACTCGTAAACGTTGAAATTTGGCCCAT 302
      |||||
QY      506 GAACCATGAGCACTTCAAAAGCGGGCGAAGTGAATTTGTTGGCATTAAGACATCGAG 565
      |||||
Db      303 GAACCTGATGAACACTTTTAAAGCGCGTGAAGTGAATTTGCTGTGTATTAAGACATTTGAG 362
      |||||
QY      566 CTCAATCAGCGCAAGAAAGTAAACCGAACTTCTGAGGTACTAATGAAGGCGCTTAAATCTCTTG 625
      |||||
Db      363 CTCAACCTGGGTAAAGGAAGTTTACCGAGCTTGAAGTTTACCAAGAAAGGACCTTAAATCTCTTG 422
      |||||
QY      626 CATGTGGGTAGCCACTTTCCACTTCTTTGAAGCTAACCAAGGCACTTAAATTCGATCGTGA 685
      |||||
Db      423 CATGTGGGTAGCCACTTTCCACTTCTTTGAACACCAAGGCACTTGAATTCGATCGGAA 482
      |||||
QY      686 AAAGCTTATGGCAAAACGCTTAGATATTCCCTCTCTGGCAACACGCTACGCAATTTGGGGCAGGA 745
      |||||
Db      483 AAAGCTTATGGCAAAACGCTTAGATATTCCCTCTCTGGCAACACGCTACGCAATTTGGGGCAGGA 542
      |||||
QY      746 CAAACCCGCAAGTGCAAGTTCCTCTCTGGTGCAGTAAATAAGTGAATTTGGCATGAAC 805
      |||||
Db      543 CAAACCCGCAAGTGCAAGTTCCTCTCTGGGCTAGTAAATAAGTGAATTTGGCATGAAC 602
      |||||
QY      806 GGGCTTGTGAATAACATCATCGCGGATGAACGCGCATAAACATAAAGCGCTTTGACAAGGCGAAA 865
      |||||
Db      603 GGGCTTGTGAATAATAATTGCGGACGAGCCATTAACACAAAGCACTAGACAAAGCAAAA 662
      |||||
QY      866 TCTCAGCGATTTCAAGTAAAGGAGCTCCCATGAAATTAAGAAATGAAGAAATATGTAAA 925
      |||||
Db      663 TCTCAGCGATTTCAAGTAAAGGAGCTCCCATGAAATTAAGAAATGAAGAAATATGTAAA 722
      |||||
QY      926 TACCTACGACCCCAACCAAGCGGATTAAGTGGCTTAGGAGATACCGATCTTTGGGCAGA 985
      |||||
Db      723 CACCTACGACCCCAACCAAGCGGATTAAGTGGCTTAGGAGATACCGATCTTTGGGCAGA 782
      |||||
QY      986 AGTAGAATCATGACTATACCACTATGGCGAAGAACTTAAATTTGGCGGGTAAAACTAT 1045
      |||||
Db      783 AGTAGAATCATGACTATACCACTATGGCGAAGAGCTCAAAATTTGGCGGGTAAAACTAT 842
      |||||
QY      1046 CCGTGAGGGTATGGGTACAGCAATAGCCCTGATGAAAAACACCCCTAGATTTAGTCAATCAC 1105
      |||||
Db      843 CCGTGAGGGTATGGGTACAGCAATAGCCCAATAGCCCAATAGCAACCTTAGATTTAGTCAATCAC 902
      |||||
QY      1106 TAACCGCATGATTTATCGACTACACCGGATTTACAAAGCGGACATTTGGATTTAAACCG 1165
      |||||
Db      903 CAACCGCATGATTTATCGACTACACCGGATTTATAAAGCCGACATTTGGTATTTAAATAATGG 962
      |||||
QY      1166 CAAATCCATGGCATTTGCAAGGAGGAGAAACAAAGGACATGCAAGATGGCGTAAAGCCCTCA 1225
      |||||
Db      963 CAAATCCATGGTATTTGGCAAGCGGGGAAACAAAGACATGCAAGATGGCGTAAAGCCCTCA 1022
      |||||
QY      1226 TATGCTCTGGGTGTGGGCACAGAACACTAGCAGGGAAGGTATGATTTATACCGCTGG 1285
      |||||
Db      1023 TATGCTCTGGGTGTGGGCACAGAACACTAGCAGGGAAGGTATGATTTATACCGCTGG 1082
      |||||
QY      1286 GGGAAATCGATTTCACACCCACTTCTCTTCTCCCAACAAATTTCCCTACCGCTCTAGCCAA 1345
      |||||
Db      1083 GGGGATCGATTTCGCAACACCCACTTCTCTCTCTCCCAACAAATTTCCCTACCGCTCTAGCCAA 1142
      |||||
QY      1346 TGGCGTTACACCACTGTTTGGAGGCGGCACAGGCTCTGTAGATGCAACCAATTCGCACTAC 1405
      |||||
Db      1143 TGGTGTTCACCACTGTTTGGAGGTGGGCACAGGCTCCGGTAGATGGCAAGTTCGCACTAC 1202
      |||||
QY      1406 TATCACTCCGGCGCAAAATGGAACTTTCACCGCATGTTTGGCGCGCAGCAGAGAGTATTCTAT 1465
      |||||
```


Db 1203 CATCACTCGGGCAAAATGAACTTGACCGCATGTTGCGCGAGCTGAAGAGTATTCTAT 1262
Qy 1466 GAATGTGGCTTTTGGGCAAAAGCAATAGCTCTAGCAAAAAACAATCTGTAGAACAAAGT 1525
Db 1263 GAATGTAGCTTTTGGGCANAGCCAAATAGTCTAGCAAAAAACAATCTGTAGAACAAAGT 1322
Qy 1526 AGAAGCGGCGGATGTTGTTTAAATGATGAAGACTGGGGCACAACCAAGTCCGAT 1585
Db 1323 AGAAGCGGCGGATGTTGTTTAAATGATGAAGACTGGGGCACAACCAAGTCCGAT 1382
Qy 1586 CGATCACTCTTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATAC 1645
Db 1383 CGATCACTCTTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATAC 1442
Qy 1646 AGTCAATGAGGAGGTTATGTAGTACACACCTTAAATGCAATGAACGCGGCGCCATCCA 1705
Db 1443 GGTCAATGAGGAGGTTATGTAGTACACACCTTAAATGCAATGAACGCGGCGCCATCCA 1502
Qy 1706 TGCCTACCACTTTGAGGAGCGGTTGGAGGACACTCACCTGATGTTTATCACCATGGCAGG 1765
Db 1503 TGCCTACCACTTTGAGGAGCGGTTGGAGGACACTCACCTGATGTTTATCACCATGGCAGG 1562
Qy 1766 CGAGCTCAATATTTACACCTCTCTCCACACCCCACTATTTCCCTATACCATTAATACGGT 1825
Db 1563 CGAGCTCAATATTTACACCTCTCTCCACACCCCACTATTTCCCTATACCATTAATACGGT 1622
Qy 1826 TGCAGAACCTTTAGACATGCTCATGATGCGACACCACTAGACAAACGATCCGCGAGGA 1885
Db 1623 TGCAGAACCTTTAGACATGCTCATGATGCGACACCACTAGACAAACGATCCGCGAGGA 1682
Qy 1886 TTTTACAATTTTCTCAAGCGGTATCGCGCCCGCTCTATCGCGCTGAAGATGTGCTCCA 1945
Db 1683 TTTTACAATTTTCTCAAGCGGTATCGCGCCCGCTCTATCGCGCTGAAGATGTGCTCCA 1742
Qy 1946 TGATATGGGTGATCGGATGACAAAGCTCGGATTCGCAAGCAATGGGCGTCGACGCGA 2005
Db 1743 TGATATGGGTGATCGGATGACAAAGCTCGGATTCGCAAGCAATGGGCGTCGACGCGA 1802
Qy 2006 AGTGATTCCTGAACTTGGCAGACTCGGATTAAGATTAAGAAATTTGGTAAAGCTTCC 2065
Db 1803 AGTGATTCCTGAACTTGGCAGACTCGGATTAAGATTAAGAAATTTGGTAAAGCTTCC 1862
Qy 2066 TGAAGATGCAAGATAAAGATAATTTTCGCTATTAAGCGCTACATCTCCAAATACACTAT 2125
Db 1863 TGAAGATGTCAGATAAAGATAAATTTTCGCTATTAAGCGCTACATCTCCAAATACACTAT 1922
Qy 2126 CAACCCCGCTTTGACCCAGCGGTGAGCGAGTATATCGGCTCTGTGGAGAGGGCAAGAT 2185
Db 1923 TAATCCCGCTTTGACCCAGCGGTGAGCGAGTATATCGGCTCTGTGGAGAGGGCAAGAT 1982
Qy 2186 CGCGACTTGGTGGTGGAAATCTCGCTTTTGGCGGTAAACCCCAAAATCGTGATCAA 2245
Db 1983 CGCGACTTGGTGGTGGAAATCTCGCTTTTGGCGGTAAACCCCAAAATCGTGATCAA 2042
Qy 2246 AGCGGTATGGTGGTCTCTGGAATGGCGGATTTCTAAGCGGTCTGTGCCACTCCCA 2305
Db 2043 AGTGCGCATGGTGGTCTCTGGAATGGCGGATTTCTAAGCGGTCTGTGCCACTCCCA 2102
Qy 2306 ACCGGTTTATTACCGGAAATGTTGGGCATACCGCAAGCGGCAAAATTTGACACCAAGCAT 2365
Db 2103 GCGGTTTATTACCGGAAATGTTGGGCATACCGCAAGCGGCAAAATTTGACACCAAGCAT 2162
Qy 2366 CACTTTTGTTCCTCAAGTCG 2385
Db 2163 CACTTTTGTTCCTCAAGTCG 2182

RESULT 6
ADQ37847
ID ADQ37847 standard; DNA; 8407 BP.
XX
AC ADQ37847;
XX

DT 07-OCT-2004 (first entry)
XX H. bizzozeronii urease gene cluster, ureABIEFGH.
XX
XX Urease; urease gene cluster; urease structural gene;
XX urease accessory gene; ureABIEFGH; Helicobacter bizzozeronii infection;
XX antibacterial; gene; db.
XX Helicobacter bizzozeronii.
XX US2004142343-A1.
XX 22-JUL-2004.
XX 12-AUG-2003; 2003US-00639273.
XX 16-AUG-2002; 2002US-0404337P.
XX (CHAN/) CHANG Y.
XX (SIMP/) SIMPSON K W.
XX (ZHUJ/) ZHU J.
XX Chang Y, Simpson KW, Zhu J;
XX WPI; 2004-533502/51.
XX GENBANK; AF330621.
XX Novel isolated nucleic acid molecule having urease gene cluster, and
XX conferring on Helicobacter bizzozeronii ability to produce urease, useful
XX as vaccine for preventing disease in mammals infected by H.bizzozeronii.
XX Claim 2; SEQ ID NO 1; 40pp; English.
XX The invention relates to an isolated nucleic acid molecule conferring on
XX Helicobacter bizzozeronii an ability to produce urease, where the nucleic
XX acid molecule is a urease gene cluster comprising at least one urease
XX structural gene and at least one urease accessory gene. The nucleic acid
XX molecule is chosen from ureA, ureB, ureF, ureG, ureH and ureI. The
XX invention also relates to an isolated protein encoded by the nucleic
XX acid, a vaccine for preventing onset of disease in mammals infected by H.
XX bizzozeronii comprising a nucleic acid and a carrier, and an isolated
XX antibody or its binding portion raised against the nucleic acid. The
XX nucleic acids, proteins and antibodies are useful for vaccinating mammals
XX against onset of disease caused by infection of H. bizzozeronii, which
XX involves administering the sequences. The sequences are useful for
XX detecting H. bizzozeronii in a sample of tissue or body fluids which
XX involves providing a nucleic acid as an antigen, providing an antibody,
XX or providing a nucleotide sequence as a probe in a nucleic acid
XX hybridisation assay, contacting the sample with the antigen or the probe,
XX and detecting any reaction which indicates that H. bizzozeronii is
XX present in the sample. This sequence represents the H. bizzozeronii
XX urease gene cluster, ureABIEFGH.
SQ Sequence 8407 BP; 2264 A; 1937 C; 1969 G; 2237 T; 0 U; 0 Other;
Query Match 41.1%; Score 1184.4; DB 12; Length 8407;
Best Local Similarity 68.5%; Pred. No. 1.5e-255;
Matches 1681; Conservative 0; Mismatches 741; Indels 32; Gaps 2;
Qy 183 AAATTTAACCAAGGAGTAAATAGGTGAAACTCACACCCAAAGAGCAAGAAAGTCTTGT 242
Db 2319 AAATTTGGTAGAGAGTTAGGATGAATTAACCCCTTAAGAGCTGACAGCTCATGT 2378
Qy 243 TATATTATGCGGCGAAGTGGCTAGAAAGCCAAAGCAGAGGGCTTAAGCTCAACCAAC 302
Db 2379 TGCATTATGCGGCGAATTGGCTAAAAAAGCAAGCAAAATGGCGTTAAGCTTAAATATA 2438
Qy 303 CCGAAGCCATTGCTTACATTAGTCCCATATTATTGGACGAAGCGCGCTGGAAGAAAAA 362
Db 2439 CTGAGGCGTAGGCCCTCATCAGTGCCTCATGTGTGGAGAGAGCCGCTGCAGGTAAAAA 2498
Qy 363 CCGTTGCCAGCTTATGGAAGAGTGCATGCACTTTTTCGAAAAAGATGAAGTAATGCCCG 422

| | | | |
|----|------|--|------|
| Db | 2499 | GTGTGGCGGATTTGATGTCAAGAGAGGACGAGACCTTCTTAAAGCTGATGATGTATGCCCGG | 2558 |
| Qy | 423 | GGGTGGGTAAATATGGTTCCCGATCTAGGTGTAGAAGCCACCTTTCCTGTAGTGTACGAAAC | 482 |
| Db | 2559 | GTGTAGCCCATATGATCCACGAAGTGGGATTTGAAGCTAACTTCCCTGTATGGGACAAAC | 2618 |
| Qy | 483 | TTGTAACTGTGAATTTGGCCCATCGAACGAGATGAGACCTTCAAAGCGGGCGAAGTGAAT | 542 |
| Db | 2619 | TGGTAAACATCCATACCCCGTTTGAAGATGGTGGGCATAAATTTGGCTCCGGGTGAAGTGA | 2678 |
| Qy | 543 | TTGGTTGCGATAAAGACATCGAGCTCAATCGCGCAAGAGTAAGTAACCGAACTTGAAGTTA | 602 |
| Db | 2679 | TTTTGAAAAACGAAGACATCACTTTGAATGAGCGCAAAACAAAGCCACCTTTAGAAGTGC | 2738 |
| Qy | 603 | CTAATGAAGGCGCTAAATCCTTGCATGTGGGTAGCCATTTCCACTTCTTTTGAAGCTAAACA | 662 |
| Db | 2739 | ATAACAAGCGGATCGCCCGTGCAGTGGGCTCCCACTTCCACTTCTTTTGAAGTGAATA | 2798 |
| Qy | 663 | AGGCATTAATAATTCGATCGTGAAGAACCTTATGGCAAAACGCCCTAGATATTCCTCTGGCA | 722 |
| Db | 2799 | AGCTTTTGGGAATTTGATCGTGAAGAACCTTATGGCAAAACGCCCTAGACATTTGCTTCTGGAA | 2858 |
| Qy | 723 | ACACGCTACCATTTGGGCGAGGACAAACCGCAAGTGCAGTTGATTCCTCTTGGTGGCA | 782 |
| Db | 2859 | CCGTGTGCGCTTTGAACCCCGGTGAGAAAAAACCCGTTGGAAATTTGATTCAAATTTGGCGGTA | 2918 |
| Qy | 783 | GTAATAAAGTGAATTTGGCATGAACCGGCTTGTGAATAACATCGCGGATGAACGCCATAAAC | 842 |
| Db | 2919 | ACCAACGCATTTACGGCTTTAACTCTCTTGTGGATCGCCAAAGCCGATCTGATGGCANAA | 2978 |
| Qy | 843 | ATAAAGCGCTTGACAAAGCGGAATTTCTCACGGATTT----- | 877 |
| Db | 2979 | AACTTGTCTCTCAACCGCGCCAAAGAACATGGCTTTTGGTGTGTGAATTTGGCGTTGCGATA | 3038 |
| Qy | 878 | ----ATCAAGTAGGAGACTCCCATGAAATGAATAA--ACAAGATATGTAATACCT | 930 |
| Db | 3039 | AAAAATTAAGAAAGGACAACTCCGATGAAAAAATCTCTCGAAAAGAAATATGTTTCTATGT | 3098 |
| Qy | 931 | ACGGACCCACAAAGCGGATAAAGTGGCTTAGAGATACCGATCTTTGGCGAGAGAAGTAG | 990 |
| Db | 3099 | ATGGACCCACTACGGGCGATTAAGTGAATTTGGCGATACCGACTGATCTTAGAAGTCG | 3158 |
| Qy | 991 | AACATGACTATACCATTTATGGGGAAGAACTTAAATTTGGCGGGTAAACATTCCTG | 1050 |
| Db | 3159 | AACATGACTGCACCCTTATGGGGAAGAAATTAAGTTTGGTGGCGTAAACCATTCGCG | 3218 |
| Qy | 1051 | AGGATATGGGTACAGACAATAGCCCTGATGAATAACCCCTAGATTTAGTCACTCAACG | 1110 |
| Db | 3219 | ATGGGATGGCACAAACCAACAGCCCCAGCAGCCAGCAACTCGATCTTTGTGCTCACTAACG | 3278 |
| Qy | 1111 | CGATGATTTATGACTACACCGGATTTACAAAGCCGACATTTGGATTTAAAAACGGCAAAA | 1170 |
| Db | 3279 | CCCTGATCGTGATTTACCGCGCATTTTAAAGCCGATATTTGGCATTTAAAAATTTGGCAAAA | 3338 |
| Qy | 1171 | TCCATGGCATTTGGCAAGCGAGAAACGAAGACATGCAAGATGGCGTGAAGCCCTCATATGG | 1230 |
| Db | 3339 | TCCATGGCATTTGGCAAGCGAGCAATAAGACATGCAAGATGGCGTTTGCAACAATCTTT | 3398 |
| Qy | 1231 | TCGTGGGTGGGACAGAGACCTAGCAGGGAGGTATGATTTATACCGCTGGGGGAA | 1290 |
| Db | 3399 | CGGTGGGCCCTGTCTACTGAGGCTTTGGCGCTGAAGGGCTGATTTGTATACAGCTGGTGGGA | 3458 |
| Qy | 1291 | TCGATTCACACACCCACTTCTTCTCCACAAATTTCCCTACCGCTTAGCCCAATGGCG | 1350 |
| Db | 3459 | TTGACACCCACATCCACTTTATTTTCTCTCCCAACAAATTCGCCACAGCATTTGGCAGCGGGA | 3518 |
| Qy | 1351 | TTCAACCATGTTTGGAGGGCGGACAGGTCCTGTAGATGGCAAGATTCGCACTACTATCA | 1410 |
| Db | 3519 | TCACAAACATGATTTGGTGGGGAACAGTCCAGCTGATGGCATTAACGCGACTACCATCA | 3578 |
| Qy | 1411 | CTCCGGGCAATGAACTTGACCGCATTTTGGCGCAGCAGAGAGATTTCTATGAATG | 1470 |
| Db | 3579 | CTCCGGGGCGGTGAACTTTAAAAACCATGTCTCCGTGCTCTGGAAGAAATATGCCATGAACT | 3638 |
| Qy | 1471 | TGGGCTTTTGGCAAGGCAATAGCTCTAGCAAAAAACAACCTTTGTAGAACAAAGTAAAG | 1533 |
| Db | 3639 | TGGGCTATTTGGGTAAAGGGAATGTGTCTTATGAACCCCTCCCTGGTGCATCAACTCGAAG | 3698 |
| Qy | 1531 | CGGCGCGGATTTGGTTTAAATTTGCATGAAGACTTGGGGCAACAACCAAGTGCATCGATC | 1590 |
| Db | 3699 | CTGGAGCCATTTGGCTTTTAAATTTCCAGAAAGCTGGGGTAGCACACCTGCGAGCATCTACC | 3758 |
| Qy | 1591 | ACTGCTTGAAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATACAGTCA | 1650 |
| Db | 3759 | ATTGCTTTGAATTTGGCTGACAAATACGATGTGCAAGTGCCTATCCACACGATACCTTGA | 3818 |
| Qy | 1651 | ATGAGSCAGGTTATGTAGATGACACCTTAAATTCGAATGAACGGCGCGCCATCCATGECCT | 1710 |
| Db | 3819 | ATGAAGCGGCTGTGTGAAGACACTTTTCAAGCCATTTGCTGGCGCACTATCCACACTT | 3878 |
| Qy | 1711 | ACCAATTTAGGAGCGGGTGGAGGACACTCACTGATGTTATCACCATGGCAGGCGAGC | 1770 |
| Db | 3879 | TCCACACTGAAGGTGTGTGGCGGCACGCTCCGATGTCAATTAAGATGTCTTGGCGGAAT | 3938 |
| Qy | 1771 | TCAAATTTCTACCTCTCCACCACCCCACTATTCCCTATACCAATTAATACGTTTGCAG | 1830 |
| Db | 3939 | TTAAATCTCTCCAGCTTTCTACCAACCCCACTTCTTCCACCGTGAATACGAAGCCG | 3998 |
| Qy | 1831 | AACACTTTAGACATGCTCATGACATGCAACCACTAGACAAACGCAATCCGCGAGGATTTTAC | 1890 |
| Db | 3999 | AACACATGGACATGTTGATGGTGTGCCACCACTTGGATAAAAACATCAAAAGAAGATGTC | 4058 |
| Qy | 1891 | AATTTTCTCAAAGCGGTATCCGCCCGGCTCTATCCGCGCTGAAGATGTGCTCCATGATA | 1950 |
| Db | 4059 | AGTTTGTCTGATTTCTAGGATTTGCGCCCCCAACCACTCCGCTGAGGACAACTCCACGATA | 4118 |
| Qy | 1951 | TGGGTGTGATCGGATGACAAAGCTCGGATTTGCAAGCAATGGGCGTGCAGGCGAAGTGA | 2010 |
| Db | 4119 | TGGGGAATTTTCTATCACAGCTCTGACTCCCAAGCGATGGCCGCTGTAGGCGAGTCA | 4178 |
| Qy | 2011 | TTCTCTCGAACTTTGGCAGACTGCGGATAAGAATAAAAAAGAAATTTGGTAAAGCTTCTTGAAG | 2070 |
| Db | 4179 | TCACCCGCACTTTGGCAACAGCGGACAAAAACAAAAAGAAATTTGGTTCCTGCTGAGG | 4238 |
| Qy | 2071 | ATGGCAAGATTAACGATTAATTTCCGCAATTAAGCGCTACATCTCCTCAATACACTATCAACC | 2130 |
| Db | 4239 | AAAAAGCGCATATGACAACTTTCCGCATCAAGCGCTACATTTTCCAAATACACCATCAACC | 4298 |
| Qy | 2131 | CGCTTTTGACCAACGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGGCAAGATCGCG | 2190 |
| Db | 4299 | CGCTATTTGCACACGCACTTTCTGAATATGTGCGCTCTGTAGAAGTGGGCAAAATTCGCCG | 4358 |
| Qy | 2191 | ACTTGTGTGTGTGGAATCTCTGCTTTTGGCGCTAAAAACCCAAATCTGATCAAAAGCGG | 2250 |
| Db | 4359 | ATTTGTGTCTTTGGAGTCTCTGCTTCTTTGGCATTAACCCCAATGATCATCAAGGCG | 4418 |
| Qy | 2251 | GTATGTGTGTCTCTCTGAATTTGGCGATTTCTTAACCGCTCTGTGCCACTCTCCCAACCGG | 2310 |
| | | | |

[illegible]

| | | | |
|----------|---------------|--|------|
| Db | 1995 | GAAGAAAAAGGGGATTAACGACAACTTTCAGATCAAAACGGTACTTGTGCTCAATAATACACCATT | 2054 |
| Qy | 2127 | AACCCCGCTTTTGACCCACCGCGTACGAGTATATACGGCTCTGTGGAAGAGGGCAAGATC | 2186 |
| Db | 2055 | AACCCAGCGNATCGCTCATGGGATTAGCGAGTATGTAGGTTCAGTAGAGTGGGCAAGTG | 2114 |
| Qy | 2187 | GCCGACTTGGTGGTGTGGGAATCCCTGCTTTTGGCGGTAAACCCAAAAATCGTGATCAAA | 2246 |
| Db | 2115 | GCTGACTTGGTATTGTGGAGTCCAGCATTCCTTTGGCGTGAACCCCAACATGATCATCAAA | 2174 |
| Qy | 2247 | GGCGGTATGGTGTCTTCTCTGAAATGGGCGATTCTAAACGGTCTGTGCGCACTCCCCAA | 2306 |
| Db | 2175 | GGCGGATTCATGTGCGTTAAAGCCAAATGGGCGATGCGAAACGCTTCTATCCCTACCCCAAA | 2234 |
| Qy | 2307 | CCGGTTTATTACCGCGAAATGTTTGGGCATCACGGCAAGCGGAAATTTGACACGACGATC | 2366 |
| Db | 2235 | CCGGTTTATTACAGAGAAATGTTGCTCATCATGGTAAAGCTAAATACGATGCAACATC | 2294 |
| Qy | 2367 | ACTTTTGTTCCTTCCAAAGTCGCTTATGAAAAATGGCGGTGAAAGAAAAGCTTGGGCTTAGAGCGC | 2426 |
| Db | 2295 | ACTTTTGTGTCCTCAAGCGGCTTATGACAAAGGCATTTAAAGAGAAATTAGGACTTTGAAAGA | 2354 |
| Qy | 2427 | CAAGTTCTACCGGTCAAAAACCTGCCGTAAACATCACCAAGAAAGACTTCAAGTTCAACGCAC | 2486 |
| Db | 2355 | CAAGTGTGCGGTGAAAAAATTTGCAGAAAATATCACTAAAAAAGACATGCAATTTCAACGCAC | 2414 |
| Qy | 2487 | AAAAACGGCAAAATCACCGTCCGATCCGAAAACCTTCGAGGTCTTTGTAGATGCAAACTC | 2546 |
| Db | 2415 | ACTACTGCTCACATTTGAAGTCAATCTCTGAAACTTACCATGTGTTCTGGATGGCAAGAA | 2474 |
| Qy | 2547 | TGCACCTCTAAACCCACCTCGCAAGTGCCTCTTAGCCGACGCTTACACTTTCTTCTTAGG | 2604 |
| Db | 2475 | GTAACCTCTAAACCCAGCAATAAAGTGAGCTTGGCGCAACTCTTTAGCATTTTCTTAGG | 2532 |
| RESULT 8 | | | |
| AAQ90180 | | | |
| XX | AAQ90180 | standard; DNA; 2619 BP. | |
| XX | AC | AAQ90180; | |
| XX | DT | 25-MAR-2003 (revised) | |
| XX | DT | 02-NOV-1995 (first entry) | |
| XX | DE | Helicobacter felis urease ureA/ureB operon. | |
| XX | KW | Urease; UreA; UreB; vaccine; Helicobacter infection; UreI; | |
| XX | KW | heat shock protein; ss. | |
| XX | OS | Helicobacter felis. | |
| XX | PH | Location/Qualifiers | |
| FT | RBS | 31..37 | |
| FT | FT | /*tag= b | |
| FT | FT | /note= "ureA Shine-Dalgarno site" | |
| FT | CDS | 43..756 | |
| FT | FT | /*tag= a | |
| FT | FT | /EC_number= "3.5.1.5" | |
| FT | FT | /note= "UreA" | |
| FT | RBS | 756..759 | |
| FT | FT | /*tag= d | |
| FT | FT | /note= "ureB Shine-Dalgarno site" | |
| FT | CDS | 766..2475 | |
| FT | FT | /*tag= c | |
| FT | FT | /EC_number= "3.5.1.5" | |
| FT | FT | /note= "UreB" | |
| XX | XX | | |
| PN | W09514093-A1. | | |
| XX | XX | | |
| PD | 26-MAY-1995. | | |
| XX | PF | 19-NOV-1993; 93WO-EP003259. | |
| XX | XX | | |

Db 386 ATGAGACATTAATAAGCGCGCAAGAACCCATTAGCTTGAAGTGAATAAAG 445
Qy 612 GGCCTAAATCCTTGCACTGTGGGTAGCCATTTCACATTTCTTTGAAGCTTAAACAAGGCATAA 671
Db 446 GCGATCGTCTGTGAGGTGGGATCAATTTCCATCTTCTCGAAGTGAATAAGCTCTGG 505
Qy 672 AATTCGATCGTGAAGAGCCTATGGCAAGCCCTAGATATTCCTCTGGCAACACGCTAC 731
Db 506 ACTTCGATCGCGCAAAAGCTTTTGGCAAGCCCTAGACATTGCATCTGGAACACGCGTGC 565
Qy 732 GCATTGGGAGAGCAAAACCCGCAAGTGCAGTGTGATTCCTCTTGGTGGCAGTAAAAAAG 791
Db 566 GCTTTGAACCCGGGAGGAAAAAGTGTGAATCTATTGCATCGCGGGGAATAAGCGCA 625
Qy 792 TGAATTCGCATGAACGGCTGTGAATAACATCGCGGATGAACCCATATAAACAATAAAGCGC 851
Db 626 TCTATGGCTTTAAATCTTTGGTGGATCGCAAGCCGATGCCGATGGTAAAAAACTCGGCT 685
Qy 852 TTGACAAGGCGAAATCTCAGGATTT----- 877
Db 686 TAAACCGCTAAGAAAAAGTTTGGGTCTGTAACTCGCGTTGTGAAGCGACTAAAG 745
Qy 878 ----ATCAAGTAAAGAGACTCCCATGAAATGAAAAAACAAGAAATATGTAAATACCTPACG 933
Db 746 ATAAACAATAAGGAAAAACCATGAAAAAGATTTACGAAAAAGAAATATGTTTCTATGTATG 805
Qy 934 GACCCACCAAGCGATAAGTGCCTTAGGAGATACCGATCTTTGGGAGAGATAGAAC 993
Db 806 GTCCCACTACCGGGATCGGTGTAGACTCGGCGACACTGATTTGATCTTTAGNAGTGGAGC 865
Qy 994 ATGACTATACCACTATGGCGAAGAACTTAAATTTGGCGGGGTAAAACTATCCGTGAGG 1053
Db 866 ATGATTGCACCACTTATGTGAAGAGATCAATTTGGGGCGGTAAAACTATCCGTGATG 925
Qy 1054 GTATGGGTGAGAGCAATAGCCCTGTATGAAAAACCCCTAGATTTAGTCATCACTAAACGCGA 1113
Db 926 GGATGAGTCAAAACCAATAGCCCTAGCTCTTATGAATTAGATTTGGTGTCTCACTAAACGCC 985
Qy 1114 TGAATATCACTACACGGGATTTACAAAGCGACATTTGGGATTAATAAGCGGCAAAATCC 1173
Db 986 TCAATGTGACTATACGGGCATTTACAAAGCCGACATTTGGGATTAATAAGACGGAAGATTG 1045
Qy 1174 ATGGCAATTCGACGAGGAGAAACAAGACATGCAAGATGCGTAAGCCCTCATATGTCG 1233
Db 1046 CAGGCATTTGCAAGGAGGCAATAAGACATGCAAGATGGCGTAGATAATAATCTTTGCG 1105
Qy 1234 TGGGTGTGGGCACAGAAGCACTAGCAGGGGAAGGTATGATTTATACCGCTGGGGGAATCG 1293
Db 1106 TAGGTCTCTACAGAGGCTTTGGCAGCTGAGGGCTTGATTGTAAACGCTGTGGCATCG 1165
Qy 1294 ATTACACACCCACTTCTTTCTCCACAACAATTCCTACCGCTTAGCOAATGGCGTTA 1353
Db 1166 ATACGCATATTCACTTTATCTCTCCCAACAATACTCCTACTGCTTTTGGCAGCGGGTTA 1225
Qy 1354 CAACCATGTTTGAGCGGCGACAGGTCTGTAGATGGCAGGATGGCACTACTATCACTC 1413
Db 1226 CAACCATGATTGGAGAGGACAGGACCTTCGGATGGCAGGATGGCACCACCATCACTC 1285
Qy 1414 CGGGCAATTTGGAATTTGCAACCGCATGTTCGCGCAGCAGAGAGTATTCTATGAATGTGG 1473
Db 1286 CCGGACGCGCTAATCTAAAAAGTATGTGTGGTGCAGCGAGAAATAGCCATGATCTAG 1345
Qy 1474 GCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAACCTTGTAGAACAGTGAAGCGG 1533
Db 1346 GCTTTTGGCTAAGGGGAATGTGCTTTACGAACCCCTCTTTACCGCATCAGATTTGAAGCAG 1405
Qy 1534 GCGCGATTTGGTTTAAATTTGATGAAGACTGGGGCAACAACCAAGTGCATCGATCACT 1593
Db 1406 GGGCGATTTGGTTTAAATTTCCAGAGACTGGGGGAAGCACACCTGCGAGCTATTTCACCAC 1465
Qy 1594 GCTTGAGCGTGGCAGATGAATAGATGTCAAGTTTGTATCCACCGGATACAGTCAATG 1653

Db 1466 GCCTCAATGTCCGCGATGAATACGATGTGCAAGTGGCTATCCACACCGATACCTTTAAAG 1525
Qy 1654 AGGCAGGTTATGTAGATGACACCCCTAAATGCAATGAACGGCGGCCCATTCATGCTACC 1713
Db 1526 AGCGGGCTGTGTAGAAAGACACCCCTAGAGGCGATTGCCGGCGCACCATTCATACCTTCC 1585
Qy 1714 ACATTGAGGAGCGGGTGGAGGACACTCACCTGATGTTATCCCATGAGGCGAGCTCA 1773
Db 1586 ACATCTGAAGGGCTGGGGGTGGACACGCTCCAGATGTTATCAAAATGGCAGGGAATTTA 1645
Qy 1774 ATATTCTACCTCTCCACCAACCCCACTATTTCCTATACCATTAATACGTTTGCAGAAC 1833
Db 1646 ACATTCTACCGCTCTACTAAACCCGACCTCTCTTCCACCAAAACACTGAAGCGGAGC 1705
Qy 1834 ACTTAGACATGCTCATGACATGCCACCACTAGACAAAACGATCCGCGAGGATTTACAAT 1893
Db 1706 ACATGGACATGTTAATGGTGTGCCACCACTTTGGATAAAAGTATCAAGGAAGATGTGCA 1765
Qy 1894 TTTCTCAAGCGCTATCCGCCCGCTCTATCGGGGCTGAAGATGTGCTCCATGATGCG 1953
Db 1766 TTGCCGATTTCGAGGATTCGCCCCCAACTATCGGGGCTGAAGACCAACTCCATGCATGCG 1825
Qy 1954 GTGTGATCGCGATCACAAGCTCGGATTCGCAAGCAATCGGGGCTGCAGCGAAGTATTC 2013
Db 1826 GGATCTTTCTATCACCAGCTCCGACTCTCAGGCTATGGGACGCGTAGCGGAGGTGATCA 1885
Qy 2014 CTCGAACTTGGCAGCTGCGGATTAAGAAATAAAAAAGAAATTTGGTAAAGCTTCCTGAAAGT 2073
Db 1886 CACGCACTTGGCAGACGACAGCAAAAAACAANAAGAGTTTGGGCGCTTGAAGAGGAAA 1945
Qy 2074 GCAAGATAACGATAAATTTCCGCAATTAAGCGCTACATCTCCAAATPACACTATCAACCC 2133
Db 1946 AAGSCGATAACGACAACTTCGCGATCAAAAGCTACATCTCTAAATACACCATCAACCC 2005
Qy 2134 CTTTGAACCGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGGCAAGATCGCGACT 2193
Db 2006 GGATCGGCGATGGGATTTCTGACTATGTGGGCTCTGTGGAAAGTGGGCAAAATACGCC 2065
Qy 2194 TGGTGTGTGGAATCCCTGCTTTTGGCGTAAACCCAAATCGTGATCAAAAGGCGGTA 2253
Db 2066 TCGTGTCTTGGAGTCCGGCTTTCTTTGGCATTAAGCCCAATATGATTTAAGGGCGGAT 2125
Qy 2254 TGGTGTCTTCTCTGAAATGGCGGATTTAAACGGCTCTGTGCCCACTCCCAACCGGTTT 2313
Db 2126 TTATTGGCTCTCTCAATGGCGGATGCCAATGCTCTATTCCCAACCCCTCAGCCGCTCT 2185
Qy 2314 ATTAACCGGAAATGTTTGGGCAATCAGGCAAGGGAATTTGACACCGAGCATCACTTTTG 2373
Db 2186 ATTACCGTGAATGTTTGGCACACCATGGGAAAAACAAATTCGACACCAATATCACTTT 2245
Qy 2374 TTTCCAAAGTCGCTATGAAAATGGCGTGAAGAAAGCTGGGCTTAGAGCGCAAGTTC 2433
Db 2246 TGTCCCAAGCGGCTTACAAGGCGGAGTCAAAAGAAAGAACTAGGGCTAGATCGCGGCGAC 2305
Qy 2434 TACCGGTCAAAAACCTCGCGTAAACATCAACAAAGAGCTTCAAGTTCAACGCAAAAACCG 2493
Db 2306 CGCAGTGAAAAACCTGTGCAATATCACTAAAAGGACCTCAATTCACAGATGTGACCG 2365
Qy 2494 CAAAAATCAGCGTCGATCCGAAAAACCTTTCGAGGCTTTTGTAGATGGCAAACTCTGCACCT 2553
Db 2366 CACATATTGATGTCAACCCCTGAAACCTATAAGGTGAAGTGGTGGCAAGAGGTAACCT 2425
Qy 2554 CTAAACCCACTCGCAAGTGCCTTAGCCGAGCGCTACACTTTCTTCTAGG 2604
Db 2426 CTAAAGCAGCAGATGAATTTGAGCCTAGCGCAACTTTATAATTTGTCTTAGG 2476

RESULT 10

AAQ75319

ID AAQ75319 standard; DNA; 2619 BP.

XX

AC AAQ75319;

XX

```

DT 16-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 22-JUN-1995 (first entry)
XX
DE Urease ureA/ureB gene.
XX
KW Urease; ureA gene; ureB gene; immunogen; vaccine; diagnostic;
KW Helicobacter pylori; heat shock protein; HSP; chaparonein; pILL205; ss.
XX
OS Helicobacter felis; ATCC 49179.
XX
FH Key Location/Qualifiers
FT RBS 31..36
FT CDS /tag= a
FT /tag= b
FT /label= ureA gene
FT /note= "encodes urease A subunit"
FT RBS 756..759
FT /tag= c
FT CDS 766..2475
FT /tag= d
FT /label= ureB gene
FT /note= "encodes urease B subunit"
XX
XX WO9426901-A1.
XX
XX 24-NOV-1994.
XX
XX 19-MAY-1994;. 94WO-EP001625.
XX
XX 19-MAY-1993; 93EP-00401309.
XX 19-NOV-1993; 93WO-EP003259.
XX
XX (INSP ) INST PASTEUR.
XX (INRM ) INST NAT SANTE & RECH MEDICALE.
XX
XX Labigne A, Suerbaum S, Ferrero R, Thiberge J;
XX WPI; 1995-006797/01.
XX P-PSDB; AAR67371.
XX
XX DNA from Helicobacter pylori and Helicobacter felis - used to develop
XX prods. for detection, treatment and prevention of Helicobacter infection.
XX
XX Disclosure; Fig 3; 168pp; English.
XX
XX Vaccine compositions include the A and B subunits (given in AAR67371) of
XX H. felis urease encoded by the ureA/ureB gene (AAQ75319) region of the
XX urease gene cluster of pILL205 (NCNM 1-1355), as well as the heat shock
XX proteins HSPA (AAR67374) and HSPB (AAR67373) encoded by the urease-
XX associated HSP gene cluster region (AAQ75321) of pILL689 (NCNM 1-1356).
XX Recombinant products are expressed in Escherichia coli. (Updated on 25-
XX MAR-2003 to correct PN field.) (Updated on 16-OCT-2003 to standardise OS
XX field)
XX
XX Sequence 2619 BP; 773 A; 570 C; 664 G; 612 T; 0 U; 0 Other;
XX
XX Query Match 36.8%; Score 1061; DB 2; Length 2619;
XX Best Local Similarity 65.5%; Pred. No. 5.4e-228;
XX Matches 1605; Conservative 0; Mismatches 805; Indels 41; Gaps 2;
XX
XX 192 ACAAGAGTAATAGTGAACCTCACACCCAAAGAGCAAGAAAGTTCTTGTATATATG 251
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 29 ATAAGAGGTTTAGGATGAACCTAACCCCTAAAGAACTAGACAAAGTTAAATGCTCAATTATG 88
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 252 CGGGCGAGTGCTAGAAAGCCCAAGCAGAGCGGCTTAAGCTCAACCAACCCGAGCCCA 311
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 89 CGGGCAGATTGGCAGAAGAACGCTTGCGCGGTGGTGAACCTCAATTTACACCGGAGCGG 148
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 312 TTGCTTACATTAAGTGCCCATATATATGACGAGCGCGCGGTGAAAAAACCCTGTGCC 371
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 149 TCGCGCTCATTAGCGGGCGTGTGATGGAAAAAGCGCGTGTATATAAAGCGTGGCGG 208

```


Qy 1414 CGGGCAAAATGGAACCTTGCACCGCATGTTGGCGCGCAGCAGAAAGATATTCTATGAATGTGG 1473
Db |||||
1286 CCGGACGCGCTAAATCTTAAAAAGATGTTGGCTGAGCGCAAGAAATAGCCCATGAATCTAG 1345
Qy |||||
1474 GCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAACTTGTAGAACAGTAGTAAGCGG 1533
Db |||||
1346 GCTTTTGGGCTAAGGGGAATGTGTCCTTACGAAACCCCTCTTACGCGCATCAGATTGAAGCAG 1405
Qy |||||
1534 GCGCGATTGCTTTTAAATTTGCATGAAGACTGGGCGCAACACCAAGTGGCATGCATCACT 1593
Db |||||
1406 GGGCGATTGCTTTTAAATCCACGAAGACTGGGGAAGCACAACCTTCGAGCTATTCAACCACT 1465
Qy |||||
1594 GCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATACAGTCAATG 1653
Db |||||
1466 GCCTCAATGTCCGCGATGAATACGATGTGCAAGTGGCTATCCACACCGATACCCCTTAAG 1525
Qy |||||
1654 AGCGAGTTATGTAGATGACACCTTAATGCAATGAACGGGCGCCATCCATCCCTACC 1713
Db |||||
1526 AGCGGGCTGTGTAGAAAGACACCTTAGAGCGATTGCGGGCGCACCATCCATACCTTCC 1585
Qy |||||
1714 ACATTGAGGAGCGGGTGGAGGACACTCACCTGATGTTATCACCATGCGAGCGAGTCA 1773
Db |||||
1586 AACTGAAGGGCTGGGGGTGGACACGCTCAGATGTTATCAAAATGGCAGGGGAATTTA 1645
Qy |||||
1774 ATATTCTACCTCTCTCCACCACCCCACTATTTCCCTATACCATTAATACGTTTCGAGAAC 1833
Db |||||
1646 ACATTCTACCGCTCTACTAAACCGACCACTTCCTTTCACCAAAAACACTGAAGCCGAGC 1705
Qy |||||
1834 ACTTAGACATGCTCATGACATGCCACCACTTAGACAAACGCAATCCGGGAGGATTTACAAT 1893
Db |||||
1706 ACATGACATGTTAATGGTGTGCCACCACCTTGGATAAAGATATCAAGGAAGATGTGCAGT 1765
Qy |||||
1894 TTTCTCAAACCGTATCCGCGGCTCTATCGCGGCTGAAGATGCTCCATCATATGG 1953
Db |||||
1766 TTGCCGATTGCGAGATTGCGCCCAAACTATPCGCGCTGAAGACCAACTCATACACATGG 1825
Qy |||||
1954 GTGTGATCGCGATGACAAGCTCGGATTTCGCAAGCAATGGGCGCTGCAGGGCAAGTGATTC 2013
Db |||||
1826 GGATCTTTCTATCACAGCTCCGACTCTCAGGCTATGGAGCGGTAGGCGAGGTGATCA 1885
Qy |||||
2014 CTCGAATCTGGCAGACTGCGGATAAGAAATAAAAAAGAAATTTGTPAAGCTTCTCTGAAGATG 2073
Db |||||
1886 CACGCACTTGGCAGACAGCAGACAAAAACAAAAAGAGTTTGGCGCTTGAAGAGGAA 1945
Qy |||||
2074 GCAAGATAACGATAATTTCCGATTAAGCGCTACATCTCCAATACACTATCAACCCCG 2133
Db |||||
1946 AAGCGGATACGACAACTTCGCGATCAACCGCTACATCTCTAAATACACCATCAACCCCG 2005
Qy |||||
2134 CTTTGACCCACGCGTGAGCGAGTATATCGCTCTGTGGAAGAGGGCAAGATCGCCGACT 2193
Db |||||
2006 GGATCGGCAATGGATTTCTGACTATGTGGCTCTGTGGAAGTGGGCAAAATACGCCGACC 2065
Qy |||||
2194 TGGTGTGTGGAATCTTCGCTTTTGGCGGTAAACCCCAAAATCGTGATCAAAAGCGGTA 2253
Db |||||
2066 TCGTGTCTTGGAGTCCGGCTTCTTTGGCATTAAGCCCAATATGATTTAAGGCGGAT 2125
Qy |||||
2254 TGGTGTCTTCTGMAATGGGCGATTCTAAGCGCTGTGCGCCACTCCCAACCGGTTT 2313
Db |||||
2126 TTAATGCGCTCTCTCAAAATGGGCGATGCCAATGCGTCTATTCCACCCCTCAGCGCGTCT 2185
Qy |||||
2314 ATTACCGGAAATGTTGGGCATCAGCGAGGCGAATTTGACACGAGATCACTTTTG 2373
Db |||||
2186 ATTACCGTGAATGTTTGGACACCATGGGAAAAACAAATTCGACACCAATATCACTTTTCG 2245
Qy |||||
2374 TTTCCAAAGTCGCTATGAATAATGGCGTGAAGAAAAAGCTGGGCTTAGAGCGCCCAAGTTC 2433
Db |||||
2246 TGTCCCAAGCGGCTTACAGCGAGGGATCAAGNAGAACTAGGGCTAGATCGCGCGCAC 2305
Qy |||||
2434 TACCGGTCAAAAATCGCGTAAATCAATCAACGAAGAAGACTTCAAGTTCAACGACAAAAAGG 2493
Db |||||
2306 CGCCACTGAAAAATCTGCGCAATATCACTAAAAGGACCTCAAATTCAAATTCGATGACCG 2365
Qy |||||
2494 CAAAAATCACCGTCGATCCGAAACCTTCGAGGTCTTTGTAGATGGCAAACTCTGCACT 2553

Db 2366 CACATATTGATGTCAACCTTAAACCTATATAGGTGAAAGTGGATGCAAGAGGTAACT 2425
Qy 2554 CTAACACCCACCTCGCAAGTGCCTTAGCCAGCGGTACACTTTCTTCTTAGG 2604
Db 2426 CTAAGCAGCAGATGAATTTAGCCCTAGCCCACTTTATTAATTTGTTCTAGG 2476

RESULT 11
AEB55130
ID AEB55130 standard; DNA; 4824 BP.
XX AEB55130;
AC AEB55130;
XX
DT 22-SEP-2005 (first entry)
XX
DE Salmonella vector pHUR3 expressing helicobacter antigens, SEQ ID 19.
XX
KW Circular; pHUR3; ds; vaccine; helicobacter pylori infection; urease;
KW immune stimulation; antibacterial; Immunostimulant; antiulcer;
KW cytosolic; antinflammatory; gastrointestinal-gen.; gastritis;
KW peptic ulcer; carcinoma.
XX
OS Helicobacter pylori.
OS Salmonella typhimurium.
OS Escherichia coli.
OS Synthetic.
XX
FN WO200132014-A2.
PN 10-MAY-2001.
XX
PF 01-NOV-2000; 2000WO-US030191.
XX
PR 01-NOV-1999; 99US-00431705.
XX
PA (ORAV-) ORAVAX INC.
XX
PI Kleanthous H, Londono-Arcila P, Freeman D;
XX
DR WPI; 2001-343379/36.
DR P-PSDB; AEB55113, AEB55114, AEB55115, AEB55116, AEB55117, AEB55118,
DR AEB55119, AEB55120, AEB55121, AEB55122, AEB55123, AEB55124, AEB55125,
DR AEB55126, AEB55127, AEB55128, AEB55129, AEB55131, AEB55132, AEB55133,
DR AEB55134, AEB55135, AEB55136, AEB55137, AEB55138, AEB55139, AEB55140,
DR AEB55141.
XX
PT Inducing an immune response against Helicobacter in mammals, useful for
treating Helicobacter induced gastroduodenal diseases.
XX
PS Disclosure; SEQ ID NO 19; 63pp; English.
XX
CC The invention relates to inducing an immune response against Helicobacter
in a mammal, comprising mucosally administering to the mammal an
attenuated Salmonella vector containing a nucleic acid molecule encoding
a Helicobacter antigen, and parentally administering to the mammal a
Helicobacter antigen. Also included is an attenuated Salmonella vector
comprising a nucleic acid molecule encoding a Helicobacter antigen. The
Helicobacter antigen is a urease, a urease subunit, or its immunogenic
fragment (encoded by the ureA and ureB genes). The mammal is at risk of
having but does not have Helicobacter infection or has a Helicobacter
infection. The attenuated Salmonella vector further comprises an htrA or
nirB promoter. The vector can be used in inducing an immune response
against Helicobacter in a mammal. The vector can be used to treat
CC Helicobacter infection. The vector and the method can be used to treat
atrophic gastritis, peptic ulcer diseases (e.g. gastric or duodenal
ulcers), and gastric carcinoma. Attenuated Salmonella vector pHUR3 was
constructed expressing UreA and UreB and including an E. coli Amp^r gene.
The present sequence represents the pHUR3 plasmid. NOTE: pHUR3 is
represented by both AEB55112 and AEB55130, the peptides expressed by
pHUR3 are shown in the sequence listing to be split between AEB55112 and
AEB55130, yet Figure 4 shows all the peptides being expressed by the one

| | | | | | | | | | | |
|----|---|---|---------------------|------------|--------------|--|--|--|--|--|
| CC | sequence, therefore all encoded peptides are cross-referenced to both | | | | | | | | | |
| XX | sequences. | | | | | | | | | |
| SQ | Sequence 4824 BP; 1389 A; 1090 C; 1146 G; 1199 T; 0 U; 0 Other; | | | | | | | | | |
| | Query Match | 36.7% | Score 1058.6; | DB 5; | Length 4824; | | | | | |
| | Best Local Similarity | 65.7% | Pred. No. 2.2e-227; | | | | | | | |
| | Matches 1606; | Conservative 0; | Mismatches 799; | Indels 38; | Gaps 3; | | | | | |
| QY | 197 | GAGTAATAGGTGAAACTCACACCAAGAGCAAGAAAGTTCTTGTTATATTATGCGGCG | 256 | | | | | | | |
| DB | 77 | GAGATCTCCATGAAACTCACCCAAAGAGTTAGATAAGTTGATGCTCCACTACGCTGGA | 136 | | | | | | | |
| QY | 257 | GAAGTGGCTAGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCCGGAAGCCATTGCT | 316 | | | | | | | |
| DB | 137 | GAATTGGCTTAAAAACGCAAGAAAAAGGCATTAAAGCTTAAGCTTATGTAGAAGCAGTAGCT | 196 | | | | | | | |
| QY | 317 | TACATTAGTGCCCATATTATTGGACGAAGCGCGCTGGAAAAAACCCTGTGCCAGCTT | 376 | | | | | | | |
| DB | 197 | TTGATTAGTGCCCATATTATTGGAAGAAGCAGAGCTGCTAAAAAGACTGCGGCTGAATTG | 256 | | | | | | | |
| QY | 377 | ATGGAAGAGTGCATGCACCTTTTGA AAAAAGATGAAGTAATGCCGGGTGGGTAAATATG | 436 | | | | | | | |
| DB | 257 | ATGCAAGAAGGGCGCACTCTTTTAAACCCAGATGATGTGATGGATGCGTGGCAAGCATG | 316 | | | | | | | |
| QY | 437 | GTTCCTCGATCTAGGTGTAGAAAGCCACTTTCTCTGATGGTACGAAACTTGTAACTGTGAAT | 496 | | | | | | | |
| DB | 317 | ATCCATGAAGTGGGTATTGAAGCGATGTTTCTGTATGGGACTAAACTCGTAACCGTGCAAT | 376 | | | | | | | |
| QY | 497 | TGGGCCCATCGAACCCAGATGAGCACTTCAAAAGCGGGCGAAGTGAATTTGGTTGGGATAAA | 556 | | | | | | | |
| DB | 377 | ACCCCTATTGAGGCCAATGTTAAATAGTTTCTGGTGAGTTG---TTCTTAAAAAATGAA | 433 | | | | | | | |
| QY | 557 | GACATCGAGCTCAATCAGCGCAAGAGAGTAACCGAACTTGAGGTTACTAATGAAGGSCCT | 616 | | | | | | | |
| DB | 434 | GACATCACTATCAACGAAGGCAAAAAGCCGTTAGCGTGAAAGTTAAAAATGTTGGCGAC | 493 | | | | | | | |
| QY | 617 | AAATCCTTGATGTTGGGTAGCCATTTCACCTTCTTTGAAGCTAACAGGCACTAAAAATTC | 676 | | | | | | | |
| DB | 494 | AGACCGGTTCAAAATTCGGGCTCACCTTCCATTCTTTGAAGTGAATAGATCGCTAGACTTT | 553 | | | | | | | |
| QY | 677 | GATCGTGA AAAAGCCTATGCAAAACGCTAGATATTCCTCTGGCAACAGCTACGCATT | 736 | | | | | | | |
| DB | 554 | GACAGAGAAAAAATCTTCGGTAAACGCTTAGACATTCGAGCGGGAACGCGTAAAGATT | 613 | | | | | | | |
| QY | 737 | GGGGCAGACAAACCGCAAGTGCAGTTGATTCCTCTTGGTGGCAGTA AAAAAGTGAAT | 796 | | | | | | | |
| DB | 614 | GAGCCTGGCGAAGAAAAATCCGTAGAAATTGATTCACATTTGGCGGTAAACAGAAGAACTTT | 673 | | | | | | | |
| QY | 797 | GGCATGAACGGGCTTGTAATAACATCGCGGATGAACGCCATAAACATAAGCGCTTGAC | 856 | | | | | | | |
| DB | 674 | GGATTTAACGCATTTGGTTGATAGCAAGCAGACAAACGAAAGCAAAAANAATTTGCTTTACAC | 733 | | | | | | | |
| QY | 857 | AAGCGGAAATCTCACGGATT-----ATCAAGT | 884 | | | | | | | |
| DB | 734 | AGAGCTAAAGAGCGTGGTTTTCATGGCGCTTAAAGCGATGACAACATATGTA AAAACAATT | 793 | | | | | | | |
| QY | 885 | AAGGAGACTCCCATGAAA-----ATGAAAAACAGAAATATGTAATACTCAGGACCCACC | 941 | | | | | | | |
| DB | 794 | AAGGAGTAAGAAATGAAAAAGATTAGCAGAAAAAGATATGTTTCTATGTATGGTCTCTACT | 853 | | | | | | | |
| QY | 942 | AAAGGCGATAAGTGGCTTTAGGAGATACCGATCTTTTGGGCAGAGTAGAACAATGACTAT | 1001 | | | | | | | |
| DB | 854 | ACAGGCGATAAAGTAGAATTTGGGCGATACAGACTTGATTCGCTGAAGTAGAACAATGACTAC | 913 | | | | | | | |
| QY | 1002 | ACCACCTATGGCGAAGAACTTAAATTTTGGCGGGGTAAAACTATCCGTGAGGGTATGGGT | 1061 | | | | | | | |
| DB | 914 | ACCATTTATGCGAAGAGCTTAAATTCGGTGGCGTAAACCCCTAAGAGAAGGCATGAGC | 973 | | | | | | | |
| QY | 1062 | CAGAGCAATAGCCCTGATGAAAAACACCTAGATTTAGTCACTAATACGGGATGAATATC | 1121 | | | | | | | |
| DB | 974 | CAATCTAACCAACCCCTAGCAAAAGAGAGTTGGAATTTAATTTACTTAACCGCTTTAATCGTG | 1033 | | | | | | | |

| | | | |
|----|------|--|------|
| QY | 1122 | GACTACACCGGGATTTTACAAAGCCGACATTTGGGATTTAAAAACGGCAAAAATTCATGGCAT | 1181 |
| DB | 1034 | GATTTACACCGGTATTTTATAAGCGGATATTGGTATTTAAAGATGGCAAAATCGCTGGCAT | 1093 |
| QY | 1182 | GGCAAGCGCAAGAAACAGGACATGCAAGATGGCGCTAAGCCCTCATATGTTGTTGGTGGT | 1241 |
| DB | 1094 | GGTAAAGCGGTAAACAAAGACATGCAAGATGGCGTTTAAAAACAATCTTAGCGTAGGTCT | 1153 |
| QY | 1242 | GGCACAGAACACCTAGCAGGGGAAGGTATGATTAATACCGCTGGGGGAATCGATTCACAC | 1301 |
| DB | 1154 | GCTACTAGAGCCTTAGCCGGTGAAGTTTGATCGTAAGCGCTGGTGGTATTGACACACAC | 1213 |
| QY | 1302 | ACCACCTCTTCTTCCACAAACAATTCCTTACCCTCTAGCCCAATGGCGTTTACAACCATG | 1361 |
| DB | 1214 | ATCCACTTCAATTTCAACCCCAACAATCCCTACAGCTTTTGAAGCGGTGTAACAACCAATG | 1273 |
| QY | 1362 | TTTGGAGCGGCACAGCTCCTGTAGATGGCAAGATGCGACTACTATCACTCCCGGCAAA | 1421 |
| DB | 1274 | ATTGGTGGTGGAAACCGGTCTCTGTGTGGCACTAATGCGACTACTATCACTCCAGGAGA | 1333 |
| QY | 1422 | TGGAACCTTGCAACCGCATGTTTGGCGCGACAGAAAGATTTCTATGAATGTGGGCTTTTTG | 1481 |
| DB | 1334 | AGAAATTTAAATTTGATGCTCAGAGCGCTGAAGAAATATTCTATGAATTTAGTTTCTTG | 1393 |
| QY | 1482 | GGCAAGGCAATAGCTCTAGCAAAAAACAATTTGTAGAACAAAGTAGAAGCGGCGCGATT | 1541 |
| DB | 1394 | GCTAAAGGTAACGCTTCTTAACGATGCGAGCTTAGCCGATCAAAATTTGAAGCGGTGCGATT | 1453 |
| QY | 1542 | GGTTTTAAATTTGATGAGTGGGCAACACACCAAGTGCATCGATCTGCTTGAGC | 1601 |
| DB | 1454 | GGCTTTTCAATTCACGAAGACTTGGGCGACCACTCTCTTCTGCAATCAATCATGCGTTAGAT | 1513 |
| QY | 1602 | GTGCGAGATGAATACGATGTGCAAGTTTGTATCCACACCGGATACAGTCAATGAGGCGAGT | 1661 |
| DB | 1514 | GTTCGCGACAAATACGATGTGCAAGTGCCTATCGCCACAGACACTTTGATGAAGCGGCT | 1573 |
| QY | 1662 | TATGTAGATGACACCTTAAATGCAATGAAACGGGCGCGCCATCCATGCTTACCACATTTAG | 1721 |
| DB | 1574 | TGTGTAGAGACACTATGCGCTGCTATTGCTGGAACGCACTATGCACTTTTCCACACTGAA | 1633 |
| QY | 1722 | GGAGCGGTGAGAGACACTCACTGTATGTTATCACCATGGCAGGCGAGCTCAATATTCTA | 1781 |
| DB | 1634 | GGCGCTGGCGCGGACACGCTCTCTGATATTTAAAGTAGCGGCTGAAACACAACATTTCT | 1693 |
| QY | 1782 | CCCTCTCCACACCCCGCACTATTCCCTATACCATTAATACGGTTTGAGAGACACTTTAGAC | 1841 |
| DB | 1694 | CCCGCTTCCACTAACCCCGCCATCCCTTTACCGTGAATACAGAAGCAGACATGAGAC | 1753 |
| QY | 1842 | ATGCTCATGACATGCCACCACTAGACAAAACGCAATCCGCGAGGATTTTACAAATTTCTCAA | 1901 |
| DB | 1754 | ATGCTTATGGTGTGCCACCACTTGGATAAAAGCATTTAAAGAAGATGTTTCAGTTTCGCTGAT | 1813 |
| QY | 1902 | AGCGTATCCGCCCCGGCTCTATCGCGCTGAAAGATGTGCTCCATGATATGGGTGTGATC | 1961 |
| DB | 1814 | TCAAGGATCCGCCCTCAAAACCATTTCCGCTGAAGACACTTTTGCAATGACATGGGGATTTTC | 1873 |
| QY | 1962 | GCGATGACAGCTCGGATTCGCAAGCAATGGGCGTGCAGCGAAGTGAATTCCTCGAACT | 2021 |
| DB | 1874 | TCAATCCAGTTCTGACTCTCAAGCGATGGCGGTGGGGTGAAGTTATCACTAGAACT | 1933 |
| QY | 2022 | TGGCAGACTGGGATAGAATAAAAAAGAAATTTGGTAAAGCTTCTCTGAAGATGGCAAGAT | 2081 |
| DB | 1934 | TGGCAACAGCTGACAAAACAGAAAGAAATTTGGCCGCTTGAAGAAGAAAAAGCGAT | 1993 |
| QY | 2082 | AACGATAATTTCCGATTAAGCGCTTACATCTCCAAATACATATCAACCCCGCTTTGACC | 2141 |
| DB | 1994 | AACGACAACTTCAGGATCAAAACGCTACTTGTCTAAATACACCATTTAAACCCAGCGATCGT | 2053 |
| QY | 2142 | CACCGGTGAGCGAGTATATCGGCTCTGTGGAAGGGCAAGATCGCCGACTTTGGTGGT | 2201 |
| DB | 2054 | CATGGGATTAGCGAGTATGTAGTTTCAGTAGAAGTGGGCAAGTGGCTGACTTTGGTATTG | 2113 |
| QY | 2202 | TGGAATCTCGCTTTTTTTTTGGCGTAAAAACCCAAAAATCGTGATCAAAAGCGGTATGGTGGTC | 2261 |

Db 2114 TGGAGTCAGCATCTTTGGCGTGAACCCCAACATCATCAAGCGGATTCATGGG 2173
Qy 2262 TTCTCTGAAATGGCGATTTAAACGGCTGTGCGCCACTCCCAACGGGTTTATTACCGC 2321
Db 2174 TTAAGCCAATGGCGATGGGAACGCTTCTATCCCTACCCCAACCGGTTATTACAGA 2233
Qy 2322 GAAATGTTGGGCATCACGGCAAGCGAAATTTGACACGAGCATCTTTTGTTCGAA 2381
Db 2234 GAAATGTTGGCTCATCATGTAAGCTAAATACGATGCAACATCATCTTTGTGTCTCA 2293
Qy 2382 GTCGCTATGAAATGGCTGAAGAAAGCTGGCTTAGAGCGCCAAAGTCTTACGGTC 2441
Db 2294 GCGGCTTATGACAAGGCAATTAAGAAGAAATTAGGACTTTGAAAGACAAGTGTTCGCGTA 2353
Qy 2442 AAAAATCGCGTAAACATCACCAAGAAAGACTTCAAGTTCAACACAAAAACGCAAAATC 2501
Db 2354 AAAAATGCGAATAATCACTAAAAAGACATGCAATTCACGACTACCGCTCACTT 2413
Qy 2502 ACCGTGATCCGAAACCTTCGAGGTCTTTGTAGATGGCAAACTCTGCACCTCTAAACCC 2561
Db 2414 GAAGTCAATCTGAACTTACCATGTTACCATGTTGCTGATGGCAAGAAAGTAACTTCTAAACCA 2473
Qy 2562 ACTCGCAAGTGCCTCTAGCCAGCGCTACACTTTCTCTTAGG 2604
Db 2474 GCCAATAAAGTGAAGTGGCGCAACTCTTAGCATTTTCTTAGG 2516

RESULT 12
ID AEB55112 standard; DNA; 4824 BP.
XX AEB55112;
AC AEB55112;
XX 22-SEP-2005 (first entry)
XX Salmonella vector pHUR3 expressing helicobacter antigens, SEQ ID 1.
XX Circular; pHUR3; ds; vaccine; helicobacter pylori infection; urease;
KW immune stimulation; antibacterial; immunostimulant; antiulcer;
KW cytotetic; antiinflammatory; gastrointestinal-gen.; gastritis;
KW peptic ulcer; carcinoma.
XX Helicobacter pylori.
OS Salmonella typhimurium.
OS Escherichia coli.
OS Synthetic.
XX WO200132014-A2.
XX 10-MAY-2001.
XX 01-NOV-2000; 2000WO-US030191.
PF 01-NOV-1999; 99US-00431705.
PR (ORAV-) ORAVAX INC.
PA Kleanthous H, Londono-Arcila P, Freeman D;
XX WPI: 2001-343379/36.
DR P-PDB; AEB55113, AEB55114, AEB55115, AEB55116, AEB55117, AEB55118,
DR AEB55119, AEB55120, AEB55121, AEB55122, AEB55123, AEB55124, AEB55125,
DR AEB55126, AEB55127, AEB55128, AEB55129, AEB55131, AEB55132, AEB55133,
DR AEB55134, AEB55135, AEB55136, AEB55137, AEB55138, AEB55139, AEB55140,
DR AEB55141.
XX Inducing an immune response against Helicobacter in mammals, useful for
PT treating Helicobacter induced gastroduodenal diseases.
XX Disclosure; SEQ ID NO 1; 63pp; English.
XX The invention relates to inducing an immune response against Helicobacter

CC in a mammal, comprising mucosally administering to the mammal an
CC attenuated salmonella vector containing a nucleic acid to the molecule encoding
CC a helicobacter antigen, and parenterally administering to the mammal a
CC helicobacter antigen. Also included is an attenuated salmonella vector
CC comprising a nucleic acid molecule encoding a helicobacter antigen. The
CC helicobacter antigen is a urease, a urease subunit, or its immunogenic
CC fragment (encoded by the ureA and ureB genes). The mammal is at risk of
CC having but does not have helicobacter infection or has a helicobacter
CC infection. The attenuated salmonella vector further comprises an htrA or
CC nirB promoter. The vector can be used in inducing an immune response
CC against helicobacter in a mammal. The vector can be used to treat
CC helicobacter infection. The vector and the method can be used to treat
CC helicobacter induced gastroduodenal diseases, including acute, chronic or
CC atrophic gastritis, peptic ulcer diseases (e.g. gastric or duodenal
CC ulcers), and gastric carcinoma. Attenuated salmonella vector pHUR3 was
CC constructed expressing UreA and UreB and including an E. coli AmpR gene.
CC The present sequence represents the pHUR3 plasmid. NOTE: pHUR3 is
CC represented by both AEB55112 and AEB55130, the peptides expressed by
CC pHUR3 are shown in the sequence listing to be split between AEB55112 and
CC AEB55130, yet Figure 4 shows all the peptides being expressed by the one
CC sequence, therefore all encoded peptides are cross-referenced to both
CC sequences.
XX
SQ Sequence 4824 BP; 1389 A; 1090 C; 1146 G; 1199 T; 0 U; 0 Other;
Query Match 36.7%; Score 1058.6; DB 5; Length 4824;
Best Local Similarity 65.7%; Pred. No. 2.2e-227;
Matches 1606; Conservative 0; Mismatches 799; Indels 38; Gaps 3;
Qy 197 GAGTAATAGTGAACCTCACACCCAAAGAGCAAGAAAGTTCTTTATATATATCGGGC 256
Db 77 GAGATCTCCATGAACCTCACCCCAAAAGAGTAGTAGTTGATGCTCCACTACCTGGA 136
Qy 257 GAGTGGCTAGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCGAGCCATTGCT 316
Db 137 GAATTGGCTAAAAACCGCAAGAAAGGCAATTAAGCTTAATATGTAGAGCAGTAGCT 196
Qy 317 TACATTAGTCCCATATATTATGGACGAGCGCGCTGAAAAAACAACCGTTGCCAGCTT 376
Db 197 TTGATTAGTCCCATATATTATGGAGAGAGAGAGCTGTGTAAAGAGACTCGCGCTGAATTG 256
Qy 377 ATGGAAGAGTGCATGCACCTTTTGAAGAAAGATGAAGTAATGCCCGGGTGGTAAATAG 436
Db 257 ATGCAAGAGGGCGCACTCTTTTAAACACAGATGATGTGATGGATGGCGTGGCAAGCATG 316
Qy 437 GTTCCCGATCTAGGTGTAGAACCCACTTTCTGTGATGTACGAACTTGTACTGTGAT 496
Db 317 ATCCATGAAGTGGGTATTGAAGCGATGTTCTGATGGGACTTAAACTCGTAACCGGTGCAT 376
Qy 497 TGGCCCATCGAAACAGATGAGCACTTCAAGCGGGGAAAGTGAATTTGGTTGGCATAAA 556
Db 377 ACCCTATTGAGGCAATGGTAATATTAGTTCTCTGGTAGTTG---TTCTTAAAAAATGNA 433
Qy 557 GACATCGAGCTCAATGCGAGCAAGAAAGATACCGAACTTGAGGTTTACTAATGAAGGCT 616
Db 434 GACATCACTATCAACGAAGGCAAAAGCCGTTAGCGTGAAGTTTAAAAATCTTGGCGAC 493
Qy 617 AAATCTTGTGATGGGTAGCCATTTCACCTTTCTTTGAAGCTAACAGGCCTAAATTC 676
Db 494 AGACCGGTTCAATCGGCTCACACTTCCATTTCTTTGAAGTGAATAGATGCCCTAGACTTT 553
Qy 677 GATCGTGAAGAGCTATGGCAACCGCTAGATATTCCTCTGGCAACACGCTAGCGATT 736
Db 554 GACAGAGAAAAAATTTTCGGTAAACCGTTAGACATTTGACGCGGAGACGCGGTAAAGATT 613
Qy 737 GGGGAGAGCAAAACCCGCAAGTGCAGTTGATTCTCTTTGGTGGCAGTAAAGAGTGAATT 796
Db 614 GAGCTGCGAGAGAAATTCGTTAGATTGATTGACATTTGGCGGTAAACAGAGATCTTT 673
Qy 797 GGCATGAACGGGCTTGTGAATTAACATCGCGATGAACGCCATTAACATAAGCGCTTGAC 856
Db 674 GGATTTAACGCTGTTGTTGATAGACAAGCAGCAACGAAAGCAAAAAAATTTCTTTTACAC 733

| | | | | |
|-----------|-------------|--|--|------|
| Db | 1814 | TC | AAGGATCGGCCTCAAAACCAATTCGGCGTGAAGACACTTTTCATGACATGGGGATTTTTC | 1873 |
| Qy | 1962 | GG | ATCACAAGCTTCGGATTTCGCAAGCAATTCGGCGGTGCGAGGCGAAGTGAATTCCTCGAACT | 2021 |
| Db | 1874 | TC | ATCACCAGTCTTGACTCTC | 1933 |
| Qy | 2022 | TG | CGAGACTCGCGATTAAGAAATAAAAAAGAATTTGGTAACTTCCTGAAGATGGCAAAAGAT | 2081 |
| Db | 1934 | TG | CAAAACAGCTGACAAAAACAAGAAAGAATTTGGCCGCTTGAAAGAAGAAAAAGCGAT | 1993 |
| Qy | 2082 | AAC | GATAAATTCGGCAATTAAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACC | 2141 |
| Db | 1994 | AAC | GACAACCTTCAGGATCAACCGCTACTTGTCTAAATACACCATTAACCCAGCGATCGCT | 2053 |
| Qy | 2142 | CAC | GGGTGAGCGAGTATATCGCTCTGTGGGAAGGCGCAAGATCGCCGACTTTGGTGGTG | 2201 |
| Db | 2054 | CAT | TGGGATTAAGCGAGTATGTAGTTCCAGTAGAAGTGGGCAAGTGGCTGACTTTGGTATTG | 2113 |
| Qy | 2202 | TG | GAATCTCTGCCTTTTTTGGCGTAAACCCAAATCGTGATCAAAAGCGGTATGTGGTGC | 2261 |
| Db | 2114 | TG | GAGTCCAGCATTTCTTTGGCGTGAACCCCAACATGATCATCAAAAGGCGGATTCATTGCG | 2173 |
| Qy | 2262 | TT | CTCTGAATATGGGCGATTCTTAACGGCTCTGTGCCACTCCCGAACCGGTTTATACCGC | 2321 |
| Db | 2174 | TT | AAGCCAAATGGGCGATCGAACGCTTCTATCCCTACCCCAACACCGGTTTATTACAGA | 2233 |
| Qy | 2322 | GAA | TGTTGGGCATCACGCGCAAGGCGAAATTTGCACACGAGCATCACTTTTGTTCCTCAA | 2381 |
| Db | 2234 | GAA | TGTTGCTCATCATGTGAAGCTTAATACGATGCAAAATCACTTTTGTGCTCAA | 2293 |
| Qy | 2382 | GT | CGCCTATGAATGGCGTGAAGAAAAAGCTTGGGCTTAGAGCGCCCAAGTTCTACCGGTC | 2441 |
| Db | 2294 | GC | GGCTTATGACAAAGCATTAAGAAAGAATTAGGACTTTGAAGACAAGTGTGGCCGGTA | 2353 |
| Qy | 2442 | AAAA | CTGCGTTAAATCACCAGAAGACTTCAAGTTCAACGACAAACGGCAAAATC | 2501 |
| Db | 2354 | AAAA | ATTGCAGAATATCACTAAAAAAGACATGCATTTCAACGACACTACCGCTCACT | 2413 |
| Qy | 2502 | ACC | GTGATCGGAAACCTTCGAGGTCTTTGTAGATGGCAACTCTGCACTCTAAACCC | 2561 |
| Db | 2414 | GAA | GTCAATCTGAAATTTACCATGTGTCGTGGATGGCAAGAAGTAATCTTAAACCA | 2473 |
| Qy | 2562 | AC | CTCGAAGTGCTCTAGCCACGCGTACACTTTCTTCTAGG | 2604 |
| Db | 2474 | GCC | AATAAGTGAGCTTGGCGCACTCTTTAGCAATTTCTAGG | 2516 |
| RESULT 13 | | | | |
| AAQ12485 | | | | |
| ID | AAQ12485 | standard; DNA; 2767 BP. | | |
| XX | AC | AAQ12485; | | |
| XX | DT | 27-AUG-2003 (revised) | | |
| DT | 23-SEP-1991 | (first entry) | | |
| XX | XX | DNA encoding A and B subunits of H. pylori urease. | | |
| DE | XX | Gastritis; peptic ulceration; duodenitis; helicobacter; campylobacter; sg. | | |
| KW | XX | Helicobacter pylori. | | |
| OS | XX | Key | | |
| XX | XX | Location/Qualifiers | | |
| XX | XX | 64..717 | | |
| FT | FT | /*tag= a | | |
| FT | FT | /label= subunit A | | |
| FT | FT | /note= "66 kD" | | |
| FT | FT | 721..2400 | | |
| FT | FT | /*tag= b | | |
| FT | FT | /label= subunit B | | |
| XX | XX | /note= "31 kD" | | |

Db 1597 ACTTACCACACTGAAGCGCTGGCGGCGACAGCTCTCTGATATTTATTTAAAGTGGCGGT 1656
Qy 1767 GAGCTCAATATTTCTACCTCTCCACCCACCCCACTATTTCCCTATACCAATTAATACGGTT 1826
Db 1657 GAAACAACATCTCCCGCTTCCACTAACCCCACTATCTCCCTTACCGTGAATCAGAA 1716
Qy 1827 GCAGAACACTTAGACATGCTCATGACATGCCACCACTTAGACAAAGCGATCCGAGGAT 1886
Db 1717 GCCGAACACATGGACATGCTTATGGTGTGCCCACTTGGATATAAAGCATTAAGAAGAT 1776
Qy 1887 TTACAATTTTCTCAAGCGCTATCGCGCCGCTCTATCGCGCTGAAGATGTGCTCAT 1946
Db 1777 GTCCAGTTGCTGATTCAGAGATTGCGCTCTCAACCAATTGCGGCTGAAGACACTTTGCAT 1836
Qy 1947 GATATGGGTGTGATCGCGATGACAAAGCTCGGATTCGCAAGCAATGGGGGTGCAGCGAA 2006
Db 1837 GACATGGGATTTTCTCAATCACTAGTCTGACTCTCAAGCGATGGCGGTGTGGTGAA 1896
Qy 2007 GTGATTCCTCGAATTTGGCGAGACTGGGATAGATATAAAGAAATTTGGTAACTTCCT 2066
Db 1897 GTTATCACTAGAACTTGGCAACAGCTGACAAAAATAAAAAAGAAATTTGGCGCTTGAAA 1956
Qy 2067 GAAGATGGCAAGATAACCATATTTCCGCTATTAAGCGCTACATCTCCAATACACTATC 2126
Db 1957 GAAGAAAAAGCGATAACGACAACTTCAGGATCAAAACGCTACTTGTCTAAATACACAT 2016
Qy 2127 AACCCCGCTTTGACCCAGCGGTGAGCGATATATCGGCTCTGTGAAAGGCGCAAGATC 2186
Db 2017 AACCCAGCGATCGCTCATGGATTAGCGAGTATGTCGGTCTGTAGAAAGTGGCAAGTG 2076
Qy 2187 GCCGACTTGGTGTGGGAATCTGCTTTTGGCGTAAACCCCAAAATCGTATCAAA 2246
Db 2077 GCTGACTTGGTATTTGGAGTCCGCAATTTCTTGGTGTGAAACCCCAACATGATCATCAA 2136
Qy 2247 GCGGTATGGTGTCTTCTGAAATGGCGGATTTCAACGCTCTGTGCCACTCCCAA 2306
Db 2137 GCGGGTTCATCGCATTGATCAATGAGTGGTGAAGCGCTTCTATCCCTACCCACAA 2196
Qy 2307 CCGGTTTATTACCGCAAAATGTTGGGCATCAGCGCAAGCGAAATTTGACACCAAGCATC 2366
Db 2197 CAGTTTATTACAGAGAAATGTTGCTCATCATGTTAAGCTTAATACATGCAACATC 2256
Qy 2367 ACTTTTGTTCAAAGTCCCTATGAAATGGCGTGAAGAAAGCTGGGCTTAGAGCGC 2426
Db 2257 ACTTTTGTCTCAAGCGCTTATGACAAAGCATTAAGAGAAATTAAGGCTTGAAGA 2316
Qy 2427 CAAGTTCTACCGTCAAAAACCTGCCGTAACTCACCAGAAAGACTTCAAGTTCAACGAC 2486
Db 2317 CAAGTGTTCGCGTAAAAAATTTGCAGAAATATCACTAAAAAAGACATGCAATTTCAACGAC 2376
Qy 2487 AAAACGGCAAAATCACCGTTCGATCCGAAACCTTCGAGGTCTTTGTAGATGGCAACTC 2546
Db 2377 ACTACCGTCACTGAAGTCAATTCGAACTTACCATGTGTCTGTGATGGCAAA-GA 2435
Qy 2547 TGCACCTCTTAAACCCACCTCGCAAGTGCCTTAGCCCCAGCGGTACACTTTCTTCTAGG 2604
Db 2436 AGTAACTTAACACCAACCAATAAAGTGAGCTGGCGCAACTCTTAGCATTTTCTAGG 2493

RESULT 14
ADQ37850
ID ADQ37850 standard; DNA; 1710 BP.
XX AC ADQ37850;
XX XX
DT 07-OCT-2004 (first entry)
DE H. bizzozeronii ureB DNA.
XX Urease; urease gene cluster; urease structural gene;
KW urease accessory gene; ureB; Helicobacter bizzozeronii infection;
KW antibacterial; gene; ds.
XX XX

OS Helicobacter bizzozeronii.
XX Key Location/Qualifiers
FH CDS 1. .1710
FT /*tag= a
FT /*product= "H. bizzozeronii ureB"
XX US2004142343-A1.
PN 22-JUL-2004.
XX
XX 12-AUG-2003; 2003US-00639273.
XX 16-AUG-2002; 2002US-0404337P.
XX (CHAN/) CHANG Y.
PA (SIMP/) SIMPSON K W.
PA (ZHU/) ZHU J.
XX Chang Y, Simpson KW, Zhu J;
XX WPI: 2004-533502/51.
XX P-PSDB; ADQ37851.
XX Novel isolated nucleic acid molecule having urease gene cluster, and
XX conferring on Helicobacter bizzozeronii ability to produce urease, useful
XX as vaccine for preventing disease in mammals infected by H.bizzozeronii.
XX Claim 7; SEQ ID NO 4; 40pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule conferring on
CC Helicobacter bizzozeronii an ability to produce urease, where the nucleic
CC acid molecule is a urease gene cluster comprising at least one urease
CC structural gene and at least one urease accessory gene. The nucleic acid
CC molecule is chosen from ureA, ureB, ureF, ureG, ureH and ureI. The
CC invention also relates to an isolated protein encoded by the nucleic
CC acid, a vaccine for preventing onset of disease in mammals infected by H.
CC bizzozeronii comprising a nucleic acid and a carrier, and an isolated
CC antibody or its binding portion raised against the nucleic acid. The
CC nucleic acids, proteins and antibodies are useful for vaccinating mammals
CC against onset of disease caused by infection of H. bizzozeronii, which
CC involves administering the sequences. The sequences are useful for
CC detecting H. bizzozeronii in a sample of tissue or body fluids which
CC involves providing a nucleic acid as an antigen, providing an antibody,
CC or providing a nucleotide sequence as a probe in a nucleic acid
CC hybridisation assay, contacting the sample with the antigen or the probe,
CC and detecting any reaction which indicates that H. bizzozeronii is
CC present in the sample. This sequence represents H. bizzozeronii ureB DNA.
XX
SQ Sequence 1710 BP; 492 A; 436 C; 386 G; 396 T; 0 U; 0 Other;

Query Match 32.6%; Score 940.8; DB 12; Length 1710;
Best Local Similarity 72.0%; Pred. No. 4.8e-201;
Matches 1227; Conservative 0; Mismatches 477; Indels 0; Gaps 0;

Qy 900 AAAATGAAAAACAGAAATATGTAATACCTACGCGACCCACCAAGGCGATAAAGTGGCG 959
Db 7 AAAATCTCTGAAAAGAAATATGTTCTATGTATGACCCACTACGCGGATTAAGTGAGA 66

Qy 960 TTAGGAGATACCGATCTTTGGGCGAGAGTAGAACATGACTATACCACTATGGCGAAGAA 1019
Db 67 TTGGCGATACCGACCTGATCTTTAGAAAGTGAACATGACTGCACCCTTATGGCGAAGAA 126

Qy 1020 CTTAAATTTGGCGGGGTAAAACTATCCGTGAGGGTATGGGTAGAGCAATAGCCCTGAT 1079
Db 127 ATTAAGTTTGGTGGCGTAAAAACCATTCGCGATGGGATGGCACAACCAACAGCCCCAGC 186

Qy 1080 GAAAAACACCTAGATTTAGTCATCACTAACGCGATGATTATCGACTACACGGGATTTAC 1139
Db 187 AGCCAGAACTCGATCTTGTGCTCACTAACGCCCTGATCGTGGATTACACCGCATTTAT 246

Qy 1140 AAAGCCGACATTTGGGATTAAAAACGGCAAAATCCATGGCATTTGGCAAGGCAAGAAACAAG 1199

Db 247 AAGCGGATATTGGCAATTAAAAATGGCAAAATCCATGGCAATTGGCAAAAGCAGGCATATAA 306
Qy 1200 GACATGCAAGATGGCGTAAGCCCTCATATATGGTGGTGGTGGGCAACAGAACCACTAGCA 1259
Db 307 GACATGCAAGATGGCGTTTGCACAACTTTTGGTGGGCGCTGCTACTAGAGCTTTGGCC 366
Qy 1260 GGGGAAGGTATGATATTATCCGCTGGGGGAATCGAATTCACACCCCACTTCTCTTCCCA 1319
Db 367 GCTGAAGGGCTGATTGTTACAGCTGGTGGATTGACACCCACACATCCACTTTATTTCTCCC 426
Qy 1320 CAACTATCCCTACCGCTTAGCGCAATGGCTTACACCAATGTTTGGAGGGGCGACAGGT 1379
Db 427 CAACAAATCCCAAGCATTTGCGAGCGGATCAACACCAATGATTGGTGGGGGAACAGGT 486
Qy 1380 CTGTAGATGGCAAGTGGCTACTATCACTCCGGGCAATGGAACTTGGACCGCATG 1439
Db 487 CAGCTGATGGGACTAAGCGGACTACATCACTCCGGGGCGCTGGAACCTTAAACCATG 546
Qy 1440 TTGCGGCGCAGAGAAGTATTCTATGAATGTGGGCTTTTGGGCAAGGCAATAGCTCT 1499
Db 547 CTCGGTGCCTCTGAAGAATATGCCAATCTGGGCTATTGGGTAAAGGGAATGTCTCT 606
Qy 1500 AGCAAAAAACAATTGTAGAACAAAGTAGAAGCGGGCCGATTTGGTTTAAATTGCATGAA 1559
Db 607 TATGAACCCCTCCCTGGTCCGATCAACTCGAAGCTGGAGCCATTGGCTTTAAAAATCCAGAA 666
Qy 1560 GACTGGGGCACACACCAAGTGGATCGATCACTGCTTGGCGTGGCGAGATGAATGAT 1619
Db 667 GACTGGGGTAGCACACCTCGAGCCATCTACCATTTGCTTGAATGTGGCTGACAAATACGAT 726
Qy 1620 GTCAAGTTTGTATCCACACCCGATACAGTCAATGAGCGAGTATGTAGATGACACCCTA 1679
Db 727 GTGCAAGTGGCTATCACACCGATACCTTGAATGAAGCGGGCTGTGTGGAAGACACTTGTG 786
Qy 1680 AATGCAATGAACGGGCGCGCCATCCATGCTTACACATTCAGGAGCGGGTGGAGGACAC 1739
Db 787 CAAAGCAATTCGTGGGCGCACTATCCACACTTCCACACTGAAGGTGTGGTGGGGCGCAC 846
Qy 1740 TCACCTGATGTTATCACCATGGCAGGAGCTCAATATTCTACCTCTCCACCAACCCCC 1799
Db 847 GCTCCGGATGTCATTAAGATGTCTGGCGAATTTAACATCTCTCCAGCTTTTACCAACCC 906
Qy 1800 ACTATTTCCCTATACCATTAATACGGTTGCGAAGACACTTAGACATGCTCATGACATGCCAC 1859
Db 907 ACATTTCTTTTACCGTGAATACAGAAGCCGAACACATGACATGTTGATGGTGGCCAC 966
Qy 1860 CACTAGACAAACGATCCGCGAGGATTTACAAATTTTCTCAAAAGCCGATATCCGCCCGGC 1919
Db 967 CACTTGGATAAAACATCAAGAAGATGTCAGTTTGTCTGATTCTAGGATTCGCCCCCAA 1026
Qy 1920 TCTATCGCGCTGAAGATGTGCTCCATGATATGGGTGATTCGGGATGACAAAGCTCGGAT 1979
Db 1027 ACCATCGCGCTGAGGACAACTCCACGATATGGGATTTTCTATCACCAGCTCTGAC 1086
Qy 1980 TCGCAAGCAATGGGCGTGCAGCGGAGTATCTCTGAACTTGGCAGATGCGGATAAG 2039
Db 1087 TCCCAAGCGATGGGCGGTAGCGAGGTCAATCCCGCGACTTGGCAAAACAGCGGCAAA 1146
Qy 2040 AATAAAAAAGAAATTTGGTAGCTTCTCTGAAGATGGCAAGATAACGATAATTTCCGCATT 2099
Db 1147 AACAAAAAGAAATTTGGTGGCTTGGCTGAGGAAAGGCGGATATGACAACTTCGGATC 1206
Qy 2100 AAGCGCTACATCTCCAAATACATATCAACCCCGCTTTGACCCACGCGGTGAGCGAGTAT 2159
Db 1207 AAGCGCTACATTTCCAAATACACCATCAACCCCGCTATTGACACAGCGCAITTTCTGAATAT 1266
Qy 2160 ATCGGCTGTGGAAGGCAAGATCGCCGACTTGGTGTGGAATCTTGCCTTTT 2219
Db 1267 GTCGGCTCTGTAGAAGTGGGCAAAATTCGCCGATTTGGTGTCTTGGAGTCTCGGCTCTTT 1326
Qy 2220 GCGGTAAACCCAAATCGTGATCAAGGCGGATGTGTCTTCTCTGAAATGGGCGAT 2279
Db 1327 GGCATTAACCCCAACATGATCATCAAGGCGGATTCATCGCACTTTCTCAATGGGCGAT 1386

Qy 2280 TCTAAACGGCTGTGTCGCCACTCCCAACCCGGTTTATTTACCGGAAATGTTTGGGCAATCAC 2339
Db 1387 GCCAATGCTCTATCCCGCACTCCCAACCCCGTATTTACCGGAAATGTTTGGCCCAAT 1446
Qy 2340 GGCAGGCGGAATTTGACACCCAGCATCACTTTTGTTCCTCAAGTGCCTATGAAATGCG 2399
Db 1447 GGTAAAGCCAAATTTGACACCAATATCACTTTTGTATCCCAAGTGGCTTATGACAAACGGC 1506
Qy 2400 GTGAAGAAAAGCTGGGCTTAGAGCGCCAAAGTTCTTACCGGTCAAAAACCTGCGGTAAACATC 2459
Db 1507 ATTAAGAGAGTTGGGCTTCAAGAGTGGTTTTCAGTTTAAACCTGCGGCAACATC 1566
Qy 2460 ACCAAGAAAGACTTCAAGTTTCAACGACAAAAACGCAAAAAATCACCGTGCATCGAAAAAC 2519
Db 1567 ACCAAAAAGACCTCAAAATTCACGATGTTACCGCACATCGAAGTCAATCTCGAAAC 1626
Qy 2520 TTCGAGTCTTTGTAGATGGCAAACTGACACCTTAAACCCACCTCGCAAGTGCCTCTA 2579
Db 1627 TACAAAGTTAAAGTGGATGGCAAGAGGTTTCTTCCAAAGCAGCGGATAAAATCAGCCTA 1686
Qy 2580 GCCCAGCGCTACACTTTCTTCTAG 2603
Db 1687 GCACAACTCTACAACTTGTCTAG 1710

RESULT 15

ABA00816

ID ABA00816 standard; cDNA; 1719 BP.

XX ABA00816;

XX AC ABA00816;

XX DT 01-APR-2003 (first entry)

XX H. felis urease B coding sequence.

Gene; rat; IgG2a; light chain; E. coli; galactosidase; immunogen;
immune response; circulatory vessel; Gut Associated Lymphoid Tissue;
GALT; immune response; antigen; Salmonella; Cholera; immunity;
Helicobacter pylori; HIV; Candida; P. gingivalis; gut; parasite; toxin;
hormone; hormone receptor; cancer; ss.

XX Helicobacter felis.

XX OS Helicobacter felis.

XX PN WO200296949-A1.

XX PD 05-DEC-2002.

XX PF 24-MAY-2002; 2002WO-AU000661.

XX PR 25-MAY-2001; 2001AU-00005241.

XX PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.

XX PI McKenzie B, Boyle J, Lew A;

XX DR WPI; 2003-156789/15.

Raising an immune response in an animal by administering composition
comprising carrier and antigen bound to targeting moiety which binds to
receptor present in circulatory vessels in Gut Associated Lymphoid
tissue.

XX Disclosure; Page 27-28; 45pp; English.

This sequence encodes H. felis urease B. This sequence was used in the
method of the invention for raising an immune response in an animal. The
method comprises administering to the animal a composition comprising a
carrier and an antigen bound to a targeting moiety which binds to at
least one receptor present in circulatory vessels in Gut Associated
Lymphoid Tissue (GALT). The method is useful for raising an immune
response in an animal against antigens from Salmonella, Cholera,
Helicobacter pylori, HIV, Candida, P. gingivalis, gut parasites, gut

CC associated toxins, gut hormones, gut hormone receptors or gut associated
CC cancers. The method is useful for raising both mucosal and systemic
CC immunity against any antigen used in the composition

XX
SQ Sequence 1719 BP; 490 A; 411 C; 421 G; 397 T; 0 U; 0 Other;

| | | | | | | | | |
|-----------------------|--------------|--|---|------|--------|--------|-------|----|
| Query Match | 31.3%; | Score | 902.4; | DB | 8; | Length | 1719; | |
| Best Local Similarity | 70.6%; | Pred. No. | 2e-192; | | | | | |
| Matches 1203; | Conservative | 0; | Mismatches | 501; | Indels | 0; | Gaps | 0; |
| QY | 900 | AAATGAAAAA | CAAGAATATGTAATACTACGACCCACCAAGGCGGATAAAGTCGCG | 959 | | | | |
| Db | 15 | AAATTTACGAAAGAAATATGTTCTATGTATGGTCCCACTACCGGGGATCGTGTAGA | 74 | | | | | |
| QY | 960 | TTAGGAGATACCGATCTTTGGGCGAAGTAGAACATGACTATATACCACTATGGCGAAGAA | 1019 | | | | | |
| Db | 75 | CTCGGCGACACTGATTTTGATCTTAGAAGTGGAGCATGATTCACCACTTATGGTGAAGAG | 134 | | | | | |
| QY | 1020 | CTTAAATTTGGCGCGGTAAACATCTACCGTGAAGGTATGGGTGAGAGCAATAGCCCTGAT | 1079 | | | | | |
| Db | 135 | ATCAAAATTTGGGGCGGTAAACATCTACCGTGAATGGGATGAGTCAAAACCAATAGCCCTAGC | 194 | | | | | |
| QY | 1080 | GAAAAACACCCCTAGATTTAGTCACTCACTAAACGCGATGATTATCGACTACACCGGATTTAC | 1139 | | | | | |
| Db | 195 | TCITTATGAATTAGATTTGGTGTCTCACTTAAACGCCCTCATTTGGGACTATACGGGCATTTAC | 254 | | | | | |
| QY | 1140 | AAAGCCGACATTTGGGATTTAAAAACGGCAAAATCCATGGCATTTGGCAAGGCGAGAAACAAG | 1199 | | | | | |
| Db | 255 | AAAGCCGACATTTGGGATTTAAAGACGGCAAGATTGAGGCATTTGGCAAGGCGAGCAATAAG | 314 | | | | | |
| QY | 1200 | GACATGCAAGATGGCGTAAAGCCCTCATATGGTCTGTGGTGGGCAAGAGCACTAGCA | 1259 | | | | | |
| Db | 315 | GACATGCAAGATGGCGTAGATAATAATCTTTTGGGTAGGTCCTGCTACAGAGGCTTTGGCA | 374 | | | | | |
| QY | 1260 | GGGGAAGGTATGATTTATACCGCTGGGGGAATCGATTACACACCACTTCCTTTCTCCA | 1319 | | | | | |
| Db | 375 | GCTGAGGGCTTGATTTGTAACCGCTGGTGGCATCGATACGCATATTCACATTTATCTCTCCC | 434 | | | | | |
| QY | 1320 | CAACAAATTCCTACCGCTCTAGCCAATGGCGTTTACAAACCATGTTTGGAGCGGCACAGGT | 1379 | | | | | |
| Db | 435 | CAACAAATCCCTACTGCTTTTGGCCAGCGGGTTTACAAACCATGATTGGAGGAGGCACAGGA | 494 | | | | | |
| QY | 1380 | CCTGTAGATGGCAGCAATGGACTACTATCACTCCGGGCAAAATGGAATTTGCAACCGCATG | 1439 | | | | | |
| Db | 495 | CTGCGGATGGCAGCAATGGCACCACCATCACTCCCGAACGCGCTAATCTTAAAAAGTATG | 554 | | | | | |
| QY | 1440 | TTGCGCGCAGCAGAGATATTTCTATGAATGGGCTTTTGGGCAAGGCAATAGCTCT | 1499 | | | | | |
| Db | 555 | TTGCGTGCAGCGAAGAAATACGCCATGAATCTAGGCTTTTGGCTAAGGGGAATGTGCT | 614 | | | | | |
| QY | 1500 | AGCAAAAAACAACTTGTAGAACAAAGTAGAGCGGCGGCGATTTGTTTTAAATTCATGAA | 1559 | | | | | |
| Db | 615 | TACGAACCTCTTTTACGCGATCAGATTTGAAGCAGGGCGGATTTGGTTTAAAAATCCAGAA | 674 | | | | | |
| QY | 1560 | GACTGGGGCAACAAACCAAGTGGCATCGATCACTGCTTGAGCGTGGCAGATGAATACGAT | 1619 | | | | | |
| Db | 675 | GACTGGGGAAGCACACCTTCGAGCTATTTCACCACTTGCTCAATGTGCCGATGAATACGAT | 734 | | | | | |
| QY | 1620 | GTGCAAGTTTGTATCCACACCGATACAGTCAATGAGGCGAGGTTATGTAGATGACACCCCTA | 1679 | | | | | |
| Db | 735 | GTGCAAGTGGCTATCCACACCGATACCCCTTAAACGAGCGGGCTGTGTAGAAGACACCCCTA | 794 | | | | | |
| QY | 1680 | AATGCATGAACCGGGCGGCCATCCATGCTTACACATTTGAGGAGCGGGTGGAGGACAC | 1739 | | | | | |
| Db | 795 | GAGGCGAATTTGCGGGCGCCACCATCCATCCCTTCCACACTGAAGGGGCTGGGGGTGGACAC | 854 | | | | | |
| QY | 1740 | TCACCTGATGTTATCAACATGGCAGGCGAGCTCAATATTTCTACCTCTCTCCACACCCCC | 1799 | | | | | |
| Db | 855 | GCTCCAGATGTTATCAAAATGGCAGGGGAATTTAACTTCTACCCGCTCTACTTAAACCG | 914 | | | | | |
| QY | 1800 | ACTATTTCCCTATACCAATTAATACCGTTTGCAGAAACACTTAGACATGCTCATGACATGCCAC | 1859 | | | | | |
| Db | 915 | ACCATTTCTTTTCAACAAAAACACTGAAGCCGAGCACATGGACATGTTAATGGTGGCCAC | 974 | | | | | |

Search completed: November 29, 2005, 00:34:36

Job time : 1579 secs

| | | | | | | | | | |
|----|------|--|------|--|--|--|--|--|--|
| QY | 1860 | CACCTAGACAAACGCAATCCGCGAGGATTTACAAATTTTCTCAAAGCCGATACGCCCCCGGC | 1919 | | | | | | |
| Db | 975 | CACCTGGATAAAAGATATCAAGGAAGATGTGCAGTTTGCAGATTCGAGGATTCGCCCCCAA | 1034 | | | | | | |
| QY | 1920 | TCTATCGCGGTGAAGATGTCTCCATGATATGGGTGTATCGCGATGACAAAGCTCGAT | 1979 | | | | | | |
| Db | 1035 | ACTATCGCGCTGAAGACCAACTCCATGACATGGGGATCTTTCTATACCAAGCTCCGAC | 1094 | | | | | | |
| QY | 1980 | TGCAAGCAATGGGGGTGCGAGCGAAGTATTCCTCGAACTTGGCAGACTGCGGATAAG | 2039 | | | | | | |
| Db | 1095 | TCTCAGCTATGGGACGCGTAGGGAGGTGATACACGCACTTGGCAGACAGCAGACAA | 1154 | | | | | | |
| QY | 2040 | AATAAAAAGAAATTTGGTAAGCTTCTGTAAGTGGCAAGATAACGATAATTTCCGCAAT | 2099 | | | | | | |
| Db | 1155 | AACAAAAAGAGTTTGGGCGCTTTGAAAGAGAAAAAGCGGATACGACAACTTCGCGATC | 1214 | | | | | | |
| QY | 2100 | AGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACCCACGGGTGAGGAGTAT | 2159 | | | | | | |
| Db | 1215 | AAACGCTACATCTCTAAATACACCATCAACCCCGCATCGCGCATGGGATTTCTGACTAT | 1274 | | | | | | |
| QY | 2160 | ATCGGCTCTGTGAAGAGGCAAGATCGCGACTTGGTGGTGTGGAATCTCTGAAATCGGCGAT | 2219 | | | | | | |
| Db | 1275 | GTGGGCTCTGTGGAAGTGGCAAAATACGCCGACCTCGTGTGAGTCCGGCTTTCTTT | 1334 | | | | | | |
| QY | 2220 | GGCTAAAAACCCAAATTCGTGATCAAGGGCGGTATGGTGGTCTTCTCTGAAATGGCGAT | 2279 | | | | | | |
| Db | 1335 | GGCATTAAGCCCAATATGATTATTAGGGCGGATTTATTCGGCTCTCTCAAAATGGCGAT | 1394 | | | | | | |
| QY | 2280 | CTTAACGCGTCTGTGCCACTCCCAACCCGGTTTATTAACCGCGAAATGTTTGGGATCAC | 2339 | | | | | | |
| Db | 1395 | GCCAAATGCGTCTATTCCCAACCCCTCAGCCGCTCTATTACCGTGAATGTTTGGACACCAT | 1454 | | | | | | |
| QY | 2340 | GGCAAGGCGAAATTTGACACCAAGTCACTTTTGTTCBAAAGTCGCTATGAAATGSC | 2399 | | | | | | |
| Db | 1455 | GGAAAAACAAATTCGACACCAATATCACTTTCGTGTCBAAAGCGGCTTACAAGCAGGG | 1514 | | | | | | |
| QY | 2400 | GTCAAAAGAAAAGCTGGGCTTTAGAGCGCCAAAGTTCTACCGGTCAAAAACCTGCGCTAACATC | 2459 | | | | | | |
| Db | 1515 | ATCAAGAGAACTAGGGCTAGATCGGTGGTATTGCCAGTGAANAACCTGTGCAATATC | 1574 | | | | | | |
| QY | 2460 | ACCAAGAAAGACTTCAAGTTCAACGACAAACCGGCAAAAATCACCGTCGATCCGAAAACC | 2519 | | | | | | |
| Db | 1575 | ACTAAAAAGGACCTCAAAATTCACGATGTGACCGCACATATTGATGTCAACCTGAAAACC | 1634 | | | | | | |
| QY | 2520 | TTGAGGTCTTTGTAGATGGCAACTCTGCACCTCTAAACCCACCTCGCAAGTGCCTTA | 2579 | | | | | | |
| Db | 1635 | TATAAGGTGAAGTGGATGGCAAGAGTTAACCTCTAAAGCAGCAGATGAATTGAGCCTA | 1694 | | | | | | |
| QY | 2580 | GCCAGCGCTACACTTTTCTTCTAG | 2603 | | | | | | |
| Db | 1695 | GCACAACCTTTATTAATTTGTTCTAG | 1718 | | | | | | |

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2005, 00:08:23 ; Search time 2049 Seconds
(without alignments)
11635.245 Million cell updates/sec

Title: US-09-904-994B-1

Perfect score: 2883
Sequence: 1 kggragatttccarctt.....aaaaagtagagccacag 2883

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134699005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA Main:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|---------|-------|----------------------|
| 1 | 2880.6 | 99.9 | 2883 | 3 | US-09-904-994B-1 |
| 2 | 2228 | 77.3 | 2452 | 3 | US-09-904-994B-13 |
| 3 | 2158.6 | 74.9 | 2405 | 3 | US-09-904-994B-4 |
| 4 | 2158.6 | 74.9 | 2407 | 3 | US-09-904-994B-10 |
| 5 | 1948 | 67.6 | 2183 | 3 | US-09-904-994B-7 |
| 6 | 1184.4 | 41.1 | 8407 | 7 | US-10-639-273-1 |
| 7 | 940.8 | 32.6 | 1710 | 7 | US-10-639-273-4 |
| 8 | 902.4 | 31.3 | 1719 | 8 | US-10-476-313-10 |
| 9 | 886.4 | 30.7 | 1710 | 7 | US-10-335-977-3849 |
| 10 | 885.8 | 30.7 | 1815 | 3 | US-09-895-913A-251 |
| 11 | 884.8 | 30.7 | 1710 | 7 | US-10-282-122A-22427 |
| 12 | 880 | 30.5 | 1710 | 9 | US-10-500-447A-5 |
| 13 | 876.8 | 30.4 | 1717 | 8 | US-10-476-313-11 |
| 14 | 855.6 | 29.7 | 2385 | 3 | US-09-402-100-1 |
| 15 | 734.6 | 25.5 | 2341 | 7 | US-10-282-122A-32950 |
| 16 | 690.6 | 24.0 | 2439 | 7 | US-10-282-122A-34390 |
| 17 | 664.4 | 23.0 | 1830121 | 7 | US-10-329-670-1 |
| 18 | 664.4 | 23.0 | 1830121 | 8 | US-10-158-865-1 |
| 19 | 664.4 | 23.0 | 1830121 | 9 | US-10-981-687-1 |
| 20 | 662.8 | 23.0 | 5966 | 2 | US-08-781-986A-22 |
| 21 | 662.8 | 23.0 | 5966 | 2 | US-10-129-624-22 |
| 22 | 658.6 | 22.8 | 1698 | 7 | US-10-282-122A-33102 |
| 23 | 634.8 | 22.0 | 3319 | 9 | US-10-795-159-460 |

ALIGNMENTS

RESULT 1

US-09-904-994B-1
; Sequence 1, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degosedsequences
; CURRENT APPLICATION NUMBER: US/09/904, 994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2883
; TYPE: DNA
; ORGANISM: Helicobacter felis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (206)..(886)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (897)..(2603)
US-09-904-994B-1

Query Match 99.9%; Score 2880.6; DB 3; Length 2883;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2883; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| Qy | 1 | RGGRAGATTTTCCARCACTTCACGACATATTTGATCTGCTGCTGGGTGTAATTCR | 60 |
| Db | 1 | RGGRAGATTTTCCARCACTTCACGACATATTTGATCTGCTGCTGGGTGTAATTCR | 60 |
| Qy | 61 | ACTTGTGTAATCTATTATTAATTTTAAATTAATTTATTTATCATATATAATATTA | 120 |
| Db | 61 | ACTTGTGTAATCTATTATTAATTTTAAATTAATTTATTTATCATATATAATATTA | 120 |
| Qy | 121 | TTACTTATATAAAAAAGTTAATAAAAAAGTAAACGAATTTAGACTATAATCCCATTCCTT | 180 |
| Db | 121 | TTACTTATATAAAAAAGTTAATAAAAAAGTAAACGAATTTAGACTATAATCCCATTCCTT | 180 |
| Qy | 181 | TAAATTTAACACAGAGTAATAGGTGAACACTCACACCCAAAGCAGCAAGAGTTCTT | 240 |
| Db | 181 | TAAATTTAACACAGAGTAATAGGTGAACACTCACACCCAAAGCAGCAAGAGTTCTT | 240 |
| Qy | 241 | GTTATATTATGCGGGCGAAGTGGCTAGAAACGCAAGCAGAGGGCTTTAAAGCTCAACCA | 300 |

| | | | |
|----|------|--|------|
| Db | 241 | GTATATTATGCGGCGAAGTGCGGTAGAAAGCGCAAGCAGAGGGCTTTAAAGCTCAACCA | 300 |
| Qy | 301 | ACCCGAAGCCATTGCTTACATTAGTCCCATATTATGGAAGAGCGCGCGTGGAAAAAA | 360 |
| Db | 301 | ACCCGAAGCCATTGCTTACATTAGTCCCATATTATGGAAGAGCGCGCGTGGAAAAAA | 360 |
| Qy | 361 | AACCGTTCGCCAGCTTATGGAAGAGTGCATGCACTTTTGGAAAAAGATGAAGTAATGCC | 420 |
| Db | 361 | AACCGTTCGCCAGCTTATGGAAGAGTGCATGCACTTTTGGAAAAAGATGAAGTAATGCC | 420 |
| Qy | 421 | CGGGTGGGTAAATAGGTTCCCGATCTAGGTCTAGAGCCACTTTCCTGATGGTACGAA | 480 |
| Db | 421 | CGGGTGGGTAAATAGGTTCCCGATCTAGGTCTAGAGCCACTTTCCTGATGGTACGAA | 480 |
| Qy | 481 | ACTTGTAACTGTGAATTTGGCCCATCGAACACAGATGAGCACTTCAAAAGCGGGCGAAGTAA | 540 |
| Db | 481 | ACTTGTAACTGTGAATTTGGCCCATCGAACACAGATGAGCACTTCAAAAGCGGGCGAAGTAA | 540 |
| Qy | 541 | ATTTCGGTTGCGATAAAGACATCGAGCTCAATGCAAGCAAAAGATGAACCGAACTTGAGGT | 600 |
| Db | 541 | ATTTCGGTTGCGATAAAGACATCGAGCTCAATGCAAGCAAAAGATGAACCGAACTTGAGGT | 600 |
| Qy | 601 | TACTAATGAAGGCTTAATCTTCGATGTGGTAGCCATTTCCACTTCTTTGAAGCTAA | 660 |
| Db | 601 | TACTAATGAAGGCTTAATCTTCGATGTGGTAGCCATTTCCACTTCTTTGAAGCTAA | 660 |
| Qy | 661 | CAAGGCACTAAATTCGATCGTGAAGAGCTATGGCAAGCGCTAGATATTCCCTCTGG | 720 |
| Db | 661 | CAAGGCACTAAATTCGATCGTGAAGAGCGCTATGGCAAGCGCTAGATATTCCCTCTGG | 720 |
| Qy | 721 | CAACACGCTACGCAATTTGGGCGAGGCAAAACCCGCAAAAGTGCACTTTCCTTGGTGG | 780 |
| Db | 721 | CAACACGCTACGCAATTTGGGCGAGGCAAAACCCGCAAAAGTGCACTTTCCTTGGTGG | 780 |
| Qy | 781 | CAGTAAAAAGATGATGGCATGAAACCGGCTTGTGAATTAACATCGCGGATGAACGCCATAA | 840 |
| Db | 781 | CAGTAAAAAGATGATGGCATGAAACCGGCTTGTGAATTAACATCGCGGATGAACGCCATAA | 840 |
| Qy | 841 | ACATAAAGCGCTTGCAAGCGCAATCTCAGCGATTTATCAAGTAAGGAGACTCCCATGA | 900 |
| Db | 841 | ACATAAAGCGCTTGCAAGCGCAATCTCAGCGATTTATCAAGTAAGGAGACTCCCATGA | 900 |
| Qy | 901 | AAATGAAAAAACAAGATATGTAATACCTACGACCCACCAAGCGGATAAAGTGGCT | 960 |
| Db | 901 | AAATGAAAAAACAAGATATGTAATACCTACGACCCACCAAGCGGATAAAGTGGCT | 960 |
| Qy | 961 | TAGGAGATACCGATCTTTGGGAGAGTAGAACAATGACTATACCACTATGGCGAAGAAC | 1020 |
| Db | 961 | TAGGAGATACCGATCTTTGGGAGAGTAGAACAATGACTATACCACTATGGCGAAGAAC | 1020 |
| Qy | 1021 | TTAAATTTGGCGGGTAAACTATCCGTGAGGTATGGTCAAGCAATAGCCCTGATG | 1080 |
| Db | 1021 | TTAAATTTGGCGGGTAAACTATCCGTGAGGTATGGTCAAGCAATAGCCCTGATG | 1080 |
| Qy | 1081 | AAAAACACCTAGATTAGTACATCACTAAAGCGGATGATTACGACTACACCGGATTTACA | 1140 |
| Db | 1081 | AAAAACACCTAGATTAGTACATCACTAAAGCGGATGATTACGACTACACCGGATTTACA | 1140 |
| Qy | 1141 | AAGCGGACATTTGGATTTAAAAACCGCAAAATCCATGGCATTTGCAAGGCAAGAAACAGG | 1200 |
| Db | 1141 | AAGCGGACATTTGGATTTAAAAACCGCAAAATCCATGGCATTTGCAAGGCAAGAAACAGG | 1200 |
| Qy | 1201 | ACATGCMAGATGGCGTAAGCCCTCATATGGTCTGGGTATGGTCAAGAGCACTAGCAG | 1260 |
| Db | 1201 | ACATGCMAGATGGCGTAAGCCCTCATATGGTCTGGGTATGGTCAAGAGCACTAGCAG | 1260 |
| Qy | 1261 | GGGAAGGTATGATTATTACCGCTGGGGGAATCGATTACACACCCACTTCTTTCTCCAC | 1320 |
| Db | 1261 | GGGAAGGTATGATTATTACCGCTGGGGGAATCGATTACACACCCACTTCTTTCTCCAC | 1320 |
| Qy | 1321 | AACAAATCCCTACCGCTCTAGCCAAATGGCGGTACAAACCTATTTGGAGCGGCACAGGTC | 1380 |
| Db | 1321 | AACAAATCCCTACCGCTCTAGCCAAATGGCGGTACAAACCTATTTGGAGCGGCACAGGTC | 1380 |
| Qy | 1381 | CTGTAGATGCGACGAATGCGACTACTATCACTCGGGCAAAATGGAACCTTGACCGCATGT | 1440 |
| Db | 1381 | CTGTAGATGCGACGAATGCGACTACTATCACTCGGGCAAAATGGAACCTTGACCGCATGT | 1440 |
| Qy | 1441 | TGCGCGCAGCAGAAGAGTATTCTATGAATGTGGCTTTTGGGCAAAAGCAATAGCTCTTA | 1500 |
| Db | 1441 | TGCGCGCAGCAGAAGAGTATTCTATGAATGTGGCTTTTGGGCAAAAGCAATAGCTCTTA | 1500 |
| Qy | 1501 | GCAAAAAACAACCTTTAGAAACAAGTAGAGCGGCGATTTGGTTTAAATTTGCATGAAG | 1560 |
| Db | 1501 | GCAAAAAACAACCTTTAGAAACAAGTAGAGCGGCGATTTGGTTTAAATTTGCATGAAG | 1560 |
| Qy | 1561 | ACTGGGGCAACAACCAAGTGCATCGATCACTGCTTTGAGCGTGGCAGATGAATACGATG | 1620 |
| Db | 1561 | ACTGGGGCAACAACCAAGTGCATCGATCACTGCTTTGAGCGTGGCAGATGAATACGATG | 1620 |
| Qy | 1621 | TGCAAGTTTGTATCCACCCGATACAGTCAATGAGGCGAGTTATGTAGATGACACCTTAA | 1680 |
| Db | 1621 | TGCAAGTTTGTATCCACCCGATACAGTCAATGAGGCGAGTTATGTAGATGACACCTTAA | 1680 |
| Qy | 1681 | ATGCAATGAACGGCGGCCATCCATGCTTACCACTTACCACTTGGAGGAGCGGTGAGGACT | 1740 |
| Db | 1681 | ATGCAATGAACGGCGGCCATCCATGCTTACCACTTACCACTTGGAGGAGCGGTGAGGACT | 1740 |
| Qy | 1741 | CACCTGATGTTATCACCATGGCAGGAGCTCAATATTTACCCCTCTCCACCCACCCCA | 1800 |
| Db | 1741 | CACCTGATGTTATCACCATGGCAGGAGCTCAATATTTACCCCTCTCCACCCACCCCA | 1800 |
| Qy | 1801 | CTATTCCTTATACATTAATACCGTTGCGAGAACCACTTAGACATGCTATGACATGCCACC | 1860 |
| Db | 1801 | CTATTCCTTATACATTAATACCGTTGCGAGAACCACTTAGACATGCTATGACATGCCACC | 1860 |
| Qy | 1861 | ACCTAGACAAACCGCATCCGCGAGGATTTACAATTTCTCAAAGCGGATCCGCCCGGCT | 1920 |
| Db | 1861 | ACCTAGACAAACCGCATCCGCGAGGATTTACAATTTCTCAAAGCGGATCCGCCCGGCT | 1920 |
| Qy | 1921 | CTATTCGGCTGAAGATGCTCCATGATATGGGTGATCGCGATGACAAAGCTCGGATTT | 1980 |
| Db | 1921 | CTATTCGGCTGAAGATGCTCCATGATATGGGTGATCGCGATGACAAAGCTCGGATTT | 1980 |
| Qy | 1981 | CGAAGCAATTTGGCGGTGCGAGGCAAGTATCTCGAACTTTGGCAGACTGCGGATAAGA | 2040 |
| Db | 1981 | CGAAGCAATTTGGCGGTGCGAGGCAAGTATCTCGAACTTTGGCAGACTGCGGATAAGA | 2040 |
| Qy | 2041 | ATAAAAAAGATTTGGTAAAGCTTCTGAAGATGCGAAGATAACGATATTTCCGCAATTA | 2100 |
| Db | 2041 | ATAAAAAAGATTTGGTAAAGCTTCTGAAGATGCGAAGATAACGATATTTCCGCAATTA | 2100 |
| Qy | 2101 | AGCGTACATCTCCAAATACACTATCAACCCCGCTTTGACCCACCGCGCTGAGCGAGTATA | 2160 |
| Db | 2101 | AGCGTACATCTCCAAATACACTATCAACCCCGCTTTGACCCACCGCGCTGAGCGAGTATA | 2160 |
| Qy | 2161 | TCGGCTCTGTGGAAGAGGCAAGATCGCCGATTTGGTGTGTGGAATCTTGCCTTTTGTG | 2220 |
| Db | 2161 | TCGGCTCTGTGGAAGAGGCAAGATCGCCGATTTGGTGTGTGGAATCTTGCCTTTTGTG | 2220 |
| Qy | 2221 | CGGTAAACCCCAAAATCGTATCAAGCGGATGTTGTTCTCTGAAATGGGCGATT | 2280 |
| Db | 2221 | CGGTAAACCCCAAAATCGTATCAAGCGGATGTTGTTCTCTGAAATGGGCGATT | 2280 |
| Qy | 2281 | CTAACCGCTCTGTGCCACTCCCAACCGGTTTATACCGCAAAATGTTTGGGCACTACG | 2340 |
| Db | 2281 | CTAACCGCTCTGTGCCACTCCCAACCGGTTTATACCGCAAAATGTTTGGGCACTACG | 2340 |
| Qy | 2341 | GCAAGCGCAATTTGACACCCAGCATCACTTTTGTTCAAAAGTGCCTATGAAAATGGCG | 2400 |
| Db | 2341 | GCAAGCGCAATTTGACACCCAGCATCACTTTTGTTCAAAAGTGCCTATGAAAATGGCG | 2400 |
| Qy | 2401 | TGAAAGAAAAAGCTGGGCTTAGAGCGCAAGTTCACCGGTCAAAAACCTGCGCGTAAATCA | 2460 |
| Db | 2401 | TGAAAGAAAAAGCTGGGCTTAGAGCGCAAGTTCACCGGTCAAAAACCTGCGCGTAAATCA | 2460 |

QY 2461 CCAGAGAGACTTCAAGTTCAACGACAAACGGCAAAAATCAACCGTCGATCCGAAACCT 2520
DB 2461 CCAGAGAGACTTCAAGTTCAACGACAAACGGCAAAAATCAACCGTCGATCCGAAACCT 2520
QY 2521 TCGAGGTCTTTGTAGATGGCAAACTCTGCACCTCTAAACCCACTCGCAAGTGCCTCTAG 2580
DB 2521 TCGAGGTCTTTGTAGATGGCAAACTCTGCACCTCTAAACCCACTCGCAAGTGCCTCTAG 2580
QY 2581 CCAGCGCTACACTTCTCTTAGGCAAAATGCCCCCTTTTGGGGCGAGGTTATTTAGGAA 2640
DB 2581 CCAGCGCTACACTTCTCTTAGGCAAAATGCCCCCTTTTGGGGCGAGGTTATTTAGGAA 2640
QY 2641 TCTTCATCAAAACGCACTGCAATCGGTCTTGGGTGCGATCGTGCCTTTTAAACCAAC 2700
DB 2641 TCTTCATCAAAACGCACTGCAATCGGTCTTGGGTGCGATCGTGCCTTTTAAACCAAC 2700
QY 2701 TTTTCATCTTTAAGCAATGCCAATTTTAAATTAATTTCTTATAATTAATATATAT 2760
DB 2701 TTTTCATCTTTAAGCAATGCCAATTTTAAATTAATTTCTTATAATTAATATATAT 2760
QY 2761 TATGCCCTCTCATTTTAAAGGAGAAATATGCGTAGGTCTTTGGTATTTGCTATGTGGGG 2820
DB 2761 TATGCCCTCTCATTTTAAAGGAGAAATATGCGTAGGTCTTTGGTATTTGCTATGTGGGG 2820
QY 2821 TTGTTTGGTGTGGCGCAAAAGGTTATGAAACCCATCGCCTCAAAAAAGTAGAAGCCAC 2880
DB 2821 TTGTTTGGTGTGGCGCAAAAGGTTATGAAACCCATCGCCTCAAAAAAGTAGAAGCCAC 2880
QY 2881 AGG 2883
DB 2881 AGG 2883

RESULT 2

US-09-904-994b-13
; Sequence 13, Application US/09904994b
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994b
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 2452
; TYPE: DNA
; ORGANISM: Helicobacter felis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (48)..(728)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (739)..(2445)
US-09-904-994b-13

Query Match 77.3%; Score 2228; DB 3; Length 2452;

Best Local Similarity 94.3%; Pred. No. 0;

Matches 2312; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 159 AGGACTATATCCCATTCGCTTTAAATTTAAACAGAGGTAATAGGTGAAACTCACAC 218
DB 1 AGGACTATATCCCATTCGCTTTAAATTTAAACAGAGGTAATAGGTGAAACTCACAC 60
QY 219 CCAAGAGCAAGAAAGTTCTTGTTATATTATCGGCGGAAGTGGCTAGAAAGCGCAAAG 278
DB 61 CCAAGAGCAAGAAAGTTCTTGTTATATTATCGGCGGAAGTGGCTAGAAAGCGCAAAG 120
QY 279 CAGAGGGCTTAAAGCTCAACCAACCGGAGCCATTTGCTTACATTAGTGCCCATATTATGG 338
DB 121 CAGAGGGCTTAAAGCTCAACCAACCGGAGCCATTTGCTTACATTAGTGCCCATATTATGG 180

QY 339 ACGAGCGCGCGTGGAAAAAACCGTTCGCCAGCTTATGGAAGAGTGCATGCACTTTT 398
DB 181 ACGAGCGCGCGTGGAAAAAACCGTTCGCCAGCTTATGGAAGAGTGCATGCACTTTT 240
QY 399 TGAAGAAAGATGAAGTAAATGCCCCGGGTGGGTAAATATGTTTCCCAGCTAGGTGTAGAAG 458
DB 241 TGAAGAAAGATGAAGTAAATGCCCCGGGTGGGTAAATATGTTTCCCAGCTAGGTGTAGAAG 300
QY 459 CCACCTTTCCCTGTATGTAAGAACTTGTAACTGTGAATTTGGCCCCATCGAAACAGATGAGC 518
DB 301 CCACCTTTCCCTGTATGTAAGAACTTGTAACTGTGAATTTGGCCCCATCGAAACAGATGAGC 360
QY 519 ACTTCAAGCGCGGCAAGTGAATTTGGTTCGGATTAAGACATCGAGCTCAATGCAAGCA 578
DB 361 ACTTCAAGCGCGGCAAGTGAATTTGGTTCGGATTAAGACATCGAGCTCAATGCAAGCA 420
QY 579 AAGAAAGTAAACGAACTTGAAGTTACTAATGAAGGCCCTAAATCCTTTCATGTGGGTAGCC 638
DB 421 AAGAAAGTAAACGAACTTGAAGTTACTAATGAAGGCCCTAAATCCTTTCATGTGGGTAGCC 480
QY 639 ATTTCCCACTTCTTTGAAGCTAAACAAAGCACTAAATTTGATCGTGAAGAAAGCTTATGCA 698
DB 481 ATTTCCCACTTCTTTGAAGCTAAACAAAGCACTAAATTTGATCGTGAAGAAAGCTTATGCA 540
QY 699 AACGCTTAGATATTCCTCTGGCAACACGCTACGCAATTTGGGGCAGGACAAACCCGCAAG 758
DB 541 AACGCTTAGATATTCCTCTGGCAACACGCTACGCAATTTGGGGCAGGACAAACCCGCAAG 600
QY 759 TGCAGTTGATCTCTTTGGTGGCAGTAAAGAGTATTTGGCATGAAACGGGCTTGTGAATA 818
DB 601 TGCAGTTGATCTCTTTGGTGGCAGTAAAGAGTATTTGGCATGAAACGGGCTTGTGAATA 660
QY 819 ACATCGCGGATGAACGCAATAAACAAGCGCTTGACAAGCGCAAAATCTCACGGATTTA 878
DB 661 ATATTCGGACGAACGCCATTAACAAGCGCTTAGACAAGCAAAATCTCACGGATTTA 720
QY 879 TCAAGTAAAGGAGACTCCCATGAAATGAAAAACAAGAAATATGTAATACCTTACGAGCC 938
DB 721 TCAAGTAAAGGAGACTCCCATGAAATGAAAAACAAGAAATATGTAATACCTTACGAGCC 780
QY 939 ACCAAGCGGATAAAGTGGCTTAGGAGATACCGATCTTTGGGCGAAGTAGAATGAC 998
DB 781 ACCAAGCGGATAAAGTGGCTTAGGAGATACCGATCTTTGGGCGAAGTAGAATGAC 840
QY 999 TATACCACTATCGCGAAGAACTTAAATTTGGCGCGGTAAACTATCCGTGAGGCTATG 1058
DB 841 TATACCACTATCGCGAAGAACTTAAATTTGGCGCGGTAAACTATCCGTGAGGCTATG 900
QY 1059 GGTCAAGCAATAGCCCTGATGAAAAACACCTAGATTTAGTTCATCACTAAACGCGATGATT 1118
DB 901 GGTCAAGCAATAGCCCTGATGAAAAACACCTAGATTTAGTTCATCACTAAACGCGATGATT 960
QY 1119 ATCGACTACACCGGATTTTACAAAGCGCAATTTGGGATTTAAAAACGGCAAAATTCATGGC 1178
DB 961 ATTGACTACACCGGATTTTACAAAGCGCAATTTGGGATTTAAAAATGGCAAAATTCATGGC 1020
QY 1179 ATTGGCAAGCAGGAAACAAAGGATGCGGTAAAGCCCTCATATGGTTCGTGGGT 1238
DB 1021 ATTGGCAAGCAGGAAACAAAGGATGCGGTAAAGCCCTCATATGGTTCGTGGGT 1080
QY 1239 GTGGGCAAGACGACATAGCAGGGGAAAGTATGATTTATCCGCTGGGGGAATCGATTCA 1298
DB 1081 GTGGGCAAGACGACATAGCAGGGGAAAGTATGATTTATCCGCTGGGGGAATCGATTCA 1140
QY 1299 CACACCCACTTCTCTTCTTCTTCCCAACAAATTCCTACCGCTCTAGCCCAATGGCGTTACAAC 1358
DB 1141 CACACCCACTTCTCTTCTTCTTCCCAACAAATTCCTACCGCTCTAGCCCAATGGCGTTACAAC 1200
QY 1359 ATGTTTGGAGCGGCAAGGTCTGTAGTGGCAAGATGCGACTACTATCACTCCGGGC 1418
DB 1201 ATGTTTGGAGCGGCAAGGTCTGTAGTGGCAAGATGCGACTACTATCACTCCGGGC 1260
QY 1419 AATGGAACCTTGACCGCATGTTGCGCGCAGCAAGAGTATTTCTATGAAATGTGGGCTTT 1478

[illegible]

| | | | |
|---|------|---|------|
| Db | 2341 | ATCACCGTCGATCCGAAACCTTCGAGGCTTTTGTGTAGATGGCAAACTCTGCACCTCTAAA | 2400 |
| Qy | 2559 | CCCACTTCGCAAGTGCCTCTAGCCACGCTACACATTTCTTCTTAGGCACAAT | 2610 |
| Db | 2401 | CCCACTCTGAAGTGCCTTAGCCCAACGCTACACATTTCTTCTTAGGCATAAT | 2452 |
| <p>RESULT 3</p> <p>US-09-904-994B-4</p> <p>; Sequence 4, Application US/09904994B</p> <p>; Publication No. US20040005325A1</p> <p>; GENERAL INFORMATION:</p> <p>; APPLICANT: AKZO NO. US20040005325A1el N.V.</p> <p>; TITLE OF INVENTION: Helicobacter vaccine</p> <p>; FILE REFERENCE: Depoedesequities</p> <p>; CURRENT APPLICATION NUMBER: US/09/904,994B</p> <p>; CURRENT FILING DATE: 2001-07-13</p> <p>; NUMBER OF SEQ ID NOS: 21</p> <p>; SOFTWARE: Patentin Ver. 2.1</p> <p>; SEQ ID NO 4</p> <p>; LENGTH: 2405</p> <p>; TYPE: DNA</p> <p>; ORGANISM: Helicobacter felis</p> <p>; FEATURE:</p> <p>; NAME/KEY: CDS</p> <p>; LOCATION: (1)..(681)</p> <p>; FEATURE:</p> <p>; NAME/KEY: CDS</p> <p>; LOCATION: (692)..(2398)</p> <p>US-09-904-994B-4</p> | | | |
| <p>Query Match 74.9%; Score 2158.6; DB 3; Length 2405;</p> <p>Best Local Similarity 93.6%; Pred. No. 0;</p> <p>Matches 2251; Conservative 0; Mismatches 154; Indels 0; Gaps 0;</p> | | | |
| Qy | 206 | GTGAAACTCACCCCAAGAGCAAGAAAGTTCTTGTATATATTATGCGGCGGAAGTGGCT | 265 |
| Db | 1 | GTGAAACTCACCCCAAGAGCAAGAAAGTTCTTGTATATATTATGCGGCGGAAGTGGCT | 60 |
| Qy | 266 | AGAAAGCGCAAGCAGAGGCTTAAAGCTCAACCAACCGAAGCCATTGCTTACATTAGT | 325 |
| Db | 61 | AGAAAGCGCAAGCAGAGGCTTAAAGCTCAATCAACCGAAGCCATTGCTTACATTAGT | 120 |
| Qy | 326 | GCCCATATTATGACGAGGCGCGGTGGAAAAAACCCTTGCCAGCTTATGGAAGAG | 385 |
| Db | 121 | GCCCATATTATGACGAGGCGCGGTGGCAAAAACCGTTGCTGAACATTATGGAAGAA | 180 |
| Qy | 386 | TGCATGCATTTTGTAAAAAAGATGAAGTAATGCGCGGGTGGGTAAATATGGTTCCCGAT | 445 |
| Db | 181 | TGTATGCATTTTGTAAAAAAGATGAAGTATGCGCGGTGGGGAAATATGGTCCCTGAT | 240 |
| Qy | 446 | CTAGTGTAGAGCCACTTTCTCTGATGTTACGAACTGTCTTAATCTGTAATTCGCCCATC | 505 |
| Db | 241 | TTGGGCGTAGAGCCACTTTTCCCGATGGCACCACAACTCGTAACCGGTAATTCGCCCAT | 300 |
| Qy | 506 | GAACGAGATGAGCACTTCAAAGCGGGCGAAGTGAATAATTTGGTTCGATAAAGACATCGAG | 565 |
| Db | 301 | GAACCTGATGNAACATTTTAAAGCCGTGAAGTGAATTTGGCTGTGATAAAGACATTGAG | 360 |
| Qy | 566 | CTCAATGACGCAAGAAAGTAACCGAACTTTGAGGTTACTAATGAAGGGCCTAAATCCCTTG | 625 |
| Db | 361 | CTCAACGCGGTGAAGGAAGTTACCGAGCTTTGAAGTTACCAACGAAGACCTAAATCCCTTG | 420 |
| Qy | 626 | CATGTGGTAGCCATTTCCACTTTTGAAGCTAACCAAGGCATTAATAATTCGATCCGTGA | 685 |
| Db | 421 | CATGTGGTAGCCATTTCCACTTTTGAAGCAACCAAGGCATTTGAATAATTCGATCCGTGA | 480 |
| Qy | 686 | AAAGCTATGCAAAACGCCCTAGATATTTCCCTCTGGCAACACGCTACGCATTGGGGCAGGA | 745 |
| Db | 481 | AAAGCTATGCAAAACGCCCTAGATATTTCCCTCTGGCAACACGCTACGCATTGGGGCAGGA | 540 |
| Qy | 746 | CAAAACCGCAAGTGCAAGTTGATTTCTTCTTGGTGGCAGTAAAAAAGTGAATGGCATGAAC | 805 |

Db 541 CAAAACCGGTAAAGTGCAGTAAATCCCTCTTTGGCGGTAGTAAAAAGTATGGCATGAAC 600
Qy 806 GGGCTTTGTGAATAACATCGCGATGAACGCATATAAAACATAAAGCGCTTGACAAGGGCAAA 865
Db 601 GGGCTTTGTGAATAATTTGGGACGAAACGCCATAAACAACAAGCACTAGACAAGGCAAAA 660
Qy 866 TCTCAGCGATTTATCAAGTAAGAGAGACTCCCATGAAATAAGAAAAACAAGAAATATGTATAA 925
Db 661 TCTCAGCGATTCATCAAGTAAGAGAGACTCCCATGAAATAAGAAAAACAAGAGTATGTATAA 720
Qy 926 TACCTACGACCAACCAAGGGCGATAAAGTGGCGCTTAGGAGATACCGATCTTTGGGCAGA 985
Db 721 CACCTACGACCAACCAAGGGCGATAAAGTGGCGCTTAGGAGATACCGATCTTTGGGCAGA 780
Qy 986 AGTAGAATGACTATACCACTATCGCGAAGAACTTTAAATTTGGCGCGGTAAAAACTAT 1045
Db 781 AGTAGAATGACTATACCACTTTAGCGNAGAGCTCAAAATTTGGCGCGGTAAAAACTAT 840
Qy 1046 CCGTAGGGTATGGGTACAGCAATAGCCCTGATGAAAAACACCCCTAGATTTAGTCATCAC 1105
Db 841 CCGTAGGGTATGGGTACAGCAATAGCCCGATGAAAAACACCTTAGATTTAGTCATCAC 900
Qy 1106 TAAACGGATGATATACATACACCGGATTTACAAAGCCGACATTTGGGATTTAAAAACGG 1165
Db 901 CAAACGGATGATATACATACACCGGATTTTAAAGCCGACATTTGGTATTTAAAAATGG 960
Qy 1166 CAAAATCCATGGCATTGGCAAGCAGCAAAACAAGGACATGCAAGATGGCGTAAGCCCTCA 1225
Db 961 CAAAATCCATGGTATTTGGCAAGCGGGGAAACAAGACATGCAAGATGGCGTAAGCCCTCA 1020
Qy 1226 TATGGTCGTGGGTGTGGGCACAGAAGCACTAGCAGGGGAAGGTATGATTTATTACCGGTGG 1285
Db 1021 TATGGTCGTGGGTGTGGGCACAGAAGCACTAGCAGGGGAAGGTATGATTTATTACCGGTGG 1080
Qy 1286 GGAATCGATTCACACACCACTTCCTTTCTTCCACAACTATCCCTACCGCTCTAGCCAA 1345
Db 1081 GGGGATCGATTCGCACACCACTTCCTCTCTCCCAACAATTTCCCTACCGCTCTAGCCAA 1140
Qy 1346 TGGCGTTACAAACCATGTTTGGAGCGGCACAGTCTCTGATGCGACGAATCGGACTAC 1405
Db 1141 TGGTGTACAAACCATGTTTGGAGGTGGCACAGGTCCGGTAGATGGCAAGATCGCAC 1200
Qy 1406 TATCACTCCGGGCAATGGAATCTGCACCGCATGTTGCGGCAGCAGAGAAGATTTCTAT 1465
Db 1201 CATCACTCCGGCAATGGAATCTGCACCGCATGTTGCGGCAGCTGAGAAGTATCTAT 1260
Qy 1466 GAATGTGGCTTTTGGGCAAAAGGCAATAGCTCTAGCAAAAAACAATTTGTAGAACAGT 1525
Db 1261 GAATGTGGCTTTTGGGCAAAAGGCAATAGCTCCAGTAAAAACAACCTCGTAGAACAGT 1320
Qy 1526 AGAAGCGGGCGGATTTGGTTTAAATTTGCATGAAGACTGGGGCACAAACCAAGTCCGAT 1585
Db 1321 AGAAGCGGGCGGATTTGGTTTAAATTTGCATGAAGACTGGGGCACAAACCAAGTCCGAT 1380
Qy 1586 CGATCACTGCTTGAGCGTGCAGATCAATACGATGTCGAAGTTTGTATCCACACCGATAC 1645
Db 1381 CGATCACTGCTTGAGCGTAGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATAC 1440
Qy 1646 AGTCAATAGGCGAGGTTATGTAGATGACACCCCTAAATGCAATGAACGGGCGCCCATCCA 1705
Db 1441 GGTCAATAGGCGAGGTTATGTAGATGACACCCCTAAATGCGATGAACGGGCGCCCATCCA 1500
Qy 1706 TGCCTTACCAATTGAGGGAGCGGGTGGAGACACTCACTCATGATGTTATCACCATGGCAGG 1765
Db 1501 TGCCTTACCAATTGAGGGAGCGGGCGGAGGACACTCACTCATGATGTTATCACCATGGCAGG 1560
Qy 1766 CGAGCTCAATATCTACCTCTCCACACCCCACTATTCCTATACCATTAATACGGT 1825
Db 1561 CGAGCTCAATATCTACCTCTCCACACCCCACTATTCCTATACCATTAATACGGT 1620
Qy 1826 TGCAGAACACTTAGACATGCTCATGATGCCACCACTAGACAAAGCAATCGCGAGGA 1885
Db 1621 TGCAGAACACTTAGACATGCTCATGATGCCACCACTAGACAAAGCAATCGCGAGGA 1680

Qy 1886 TTTACAAATTTTCTCAAAAGCCGTATCCGCCCGGCTCTATCCGGGCTGAAGATGTGCTCCA 1945
Db 1681 TCTCCAGTTTTCCAAAGCCGTATCCGCCCGGCTCTATTCGCCGTGAAGATGTGCTCCA 1740
Qy 1946 TGATATGGGTGTGATCGCGATGACAAAGCTCCGATTTCCGAAGCAATGGGCGCTGCAGGCA 2005
Db 1741 TGATATGGGTGTGATCGCGATGACAAAGCTCCGATTTCCGAAGCAATGGGCGCGCTGGGA 1800
Qy 2006 AGTGATTTCTCGAACTTTGGCAGACTGCGGATAAGAAATAAAAAAGAAATTTGGTAAGCTTC 2065
Db 1801 AGTGATTTCTAGAACTTTGGCAAACTGCAGACAGAAATAAAAAAGAAATTTGGTAAGCTTC 1860
Qy 2066 TGAAGATGGCAAGATAACGATAATTTCCGCAATTAAGCGCTACATCTCATAATACACTAT 2125
Db 1861 TGAAGATGGTCAGATAATGACAACTTCGCGCATCAAAACGCTATATCTCCAAATACACCAT 1920
Qy 2126 CAACCCCGCTTTGACCCACGGCGTAGCGAGTATATCGGCTCTGTGGAGAGGGCAAGAT 2185
Db 1921 TAAATCCCGCTTTGACCCATGCGCGTAGCGAGTATATCGGCTCTGTGGAGAGGGCAAGAT 1980
Qy 2186 CGCGACTTGGTGTGGGAATCTGCTTTTTTTGGCGTAAAAACCAAAATCGTGATCAA 2245
Db 1981 CGCGACTTGGTGTGGGAATCTGCTTTTTCTTTGGTGTAAACCAAAATCGTGATCAA 2040
Qy 2246 AGCGGTATGGTGTGCTTTCTCTGAAAATGGGCGATTTCTAACGCGTCTGTGCCCACTCCCCA 2305
Db 2041 AGCGGTATGGTGTGTTCTCTGAAAATGGGCGATTTCTAACGCGTCTGTGCCCACTCCA 2100
Qy 2306 ACCGTTTTATTACCCGGAATGTTTGGGCATCACGGCAAGGGAAATTTGACACCAAGAT 2365
Db 2101 GCGGTTTTATTACCCGGAATGTTTGGGCATCACGGCAAGGGCAAAATTTGACACCAAGAT 2160
Qy 2366 CACTTTTGTTCGAAAGTCGCTATGAAAATGGGCGTGAAGAAAGCTGGGCTTAGAGCG 2425
Db 2161 CACTTTTGTTCGAAAGTCGCTATGAAAATGGGCGTGAAGAAAGCTTAGAGCG 2220
Qy 2426 CCAAGTTCTACCGGTCAAAATCTGCCGTAAACATCACCAAGAAAGACTTCAAGTTCAACGA 2485
Db 2221 CAAGGTCTACCGGTGAANAATCTCCGCAACATCACTAAGAAAGACTTCAAAATCAACAA 2280
Qy 2486 CAAAACGGCAAAATACCGTCGATCCGAAACCTTCGAGGTCTTTGTAGATGGCAAACT 2545
Db 2281 CAAGCGGCGCATATCACTGTGCGATCTTAAACCTTCGAGGTCTTTGTAGATGGCAAACT 2340
Qy 2546 CTGCACCTTAAACCCACCTCGAAAGTCCTCTAGCCCGAGCTACATTTTCTCTAGGC 2605
Db 2341 CTGCACCTTAAACCCCGCTCTGAAGTGCCTCTAGCCCAACGCTACACTTTTCTCTAGGC 2400
Qy 2606 ACAAT 2610
Db 2401 ACAAT 2405

RESULT 4
US-09-904-994B-10
; Sequence 10, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 2407
; TYPE: DNA
; ORGANISM: Helicobacter felis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(682)


```
Db 2042 AGCGGTATGGTGGTCTCTCTCTGAAATGGCGACTCCAAACGGCGTCCGTGGCTACACCTCA 2101
Qy 2306 ACCGGTTTATTACCGGGAATGTTTGGGCATCAGCGCAAGCGGAAATTTTACACACCAT 2365
Db 2102 GCCGGTTTATTACCGGGAATGTTTGGGCATCAGCGCAAGCGGAAATTTTACACCAT 2161
Qy 2366 CACTTTTGTGTTCCAAAGTCGCCTATGAAAAATGGCGTGAAAGAAAAGCTGGCGTTAGAGCG 2425
Db 2162 CACTTTTGTGTTCCAAAGTCGCCTATGAAAAATGGCGTGAAAGAAAAGCTAGGCTTAGAGCG 2221
Qy 2426 CCAAGTTTACCGGTCAAAAATCGCGTAACATCAACCAAGAAAGACTTCAAGTTCAACGA 2485
Db 2222 CAAAGTGCTACCGGTCAAAAATCGCGTAACATCAACCAAGAAAGACTTCAAAATTCACAA 2281
Qy 2486 CAAAACGGCAAAATCACCGTCGATCCGAAAACCTTCGAGGTCTTTGTAGATGCAAACT 2545
Db 2282 CAAAGCGGCGCATATCACTGTCATCCTAAGCTCTAGAGGTCTTTGTAGATGCAAACT 2341
Qy 2546 CTGCACCTCTAAACCCACCTCGCAAGTCCCTCTAGCCAGCGCTACACTTTTCTTAGGC 2605
Db 2342 CTGCACCTCTAAACCCGCTCTGAAAGTCCCTCTAGCCAGCGCTACACTTTTCTTAGGC 2401
Qy 2606 ACAATG 2611
Db 2402 NCAATG 2407

RESULT 5
US-09-904-994B-7
; Sequence 7, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 7
; LENGTH: 2183
; TYPE: DNA
; ORGANISM: Helicobacter felis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)..(683)
; NAME/KEY: CDS
; LOCATION: (694)..(2181)
US-09-904-994B-7

Query Match 67.6%; Score 1948; DB 3; Length 2183;
Best Local Similarity 93.3%; Pred. No. 0;
Matches 2035; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

Qy 206 GTGAATCTACACCCAAAGAGCAAGAAAGTCTTGTATATATATGCGGCGAAGTGGCT 265
Db 3 GTGAATCTACACCCAAAGAGCAAGAAAGTCTTGTATATATATGCGGCGAAGTGGCT 62
Qy 266 AGAAGCGCAAGCAGGCGCTTAAAGCTCAACCAACCCGCAAGCCATTGCTTACATTAGT 325
Db 63 AGAAGCGCAAGCAGGCGCTTAAAGCTCAATCAACCCGCAAGCCATTGCTTACATTAGT 122
Qy 326 GCCATATTATGACCAAGCGCGCGTGGAAAAAACCCTTGCCAGCTTATGGAAGAG 385
Db 123 GCCATATTATGACCAAGCGCGCGTGGAAAAAACCCTTGCTGAACTTATGGAAGAA 182
Qy 386 TGCATGCACCTTTTGAAGAAAGATGAAGTAATCCCGGGGTGGGTAAATATGGTTCCCGAT 445
Db 183 TGTATGCACCTTTTGAAGAAAGATGAGGTGATGCCCGGTGTTGGGGAATATGGTCCCTGAT 242
Qy 446 CTAGGTGTAGAACCCACTTTCCTGATGGTACGAACTTGTAACTGTGAATTGGCCCATC 505
```

```
Db 243 TTGGCGGTAGAACCACTTTCCCGATGGCACCAAACTCGTAACCGTGAATTTGGCCCAT 302
Qy 506 GAACAGATGAGCACTTTCAAAGCGGCGGAAGTGAATTTGGTTGCGATAAAGACATCGAG 565
Db 303 GAACCTGATGAACACTTTTAAAGCCGCTGAAGTGAATTTGGCTGTGATAAAGACATTGAG 362
Qy 566 CTCAATGAGGCAAAAGAGTAACCGAACTTGAGGTACTTAATGAAGGGGCTAAATCCTTTG 625
Db 363 CTCAACGTTGGGTAAGGAAGTTTACCGAGCTTGAAGTTTACCAACGAAGGACCTAAATCCTTTG 422
Qy 626 CATGTGGGTAGCACTTTCCACTTCTTTGAAGCTAAACAGGCACTAAATTTCCATCGTGAA 685
Db 423 CATGTGGGTAGCACTTTCCACTTCTTTTGAAGCCAAACAGGCACTTGAATTTCCATCGGGAA 482
Qy 686 AAAGCCTATGCAAAACGGCTAGATATTCCTCTCTGGCAACACGCTACGCACTTGGGGCAGGA 745
Db 483 AAAGCCTATGCAAAACGGCTAGATATTCCTCTCTGGCAACACGCTACGCACTTGGGGCAGGA 542
Qy 746 CAAACCCGCAAGTGCAGTTGATTCCTCTCTGGTGGCAGTAAAAAAAGTGAATGGCAATGAAC 805
Db 543 CAAACCCGTAAGTGCAGTTAATCCCTCTTGGCGGTAGTAAAAAAGTGAATGGCAATGAAC 602
Qy 806 GGGCTTGTGAATAACATCGCGGATGAACGCGCATAAACATAAAGCGCTTGACAAAGCGGAAA 865
Db 603 GGGCTTGTGAATAATAATTGCGGACGAACGCGCATAAACATAAAGCGCTTGACAAAGCGGAAA 662
Qy 866 TCTCACCGATTTATCAAGTAGGAGACTCCCATGAARATGAAAAACAAAGAAATATGTAAA 925
Db 663 TCTCACGATTTATCAAGTAGGAGACTCCCATGAARATGAAAAACAAAGAAATATGTAAA 722
Qy 926 TACCTACGAGCCACCAAGGCGATAAAGTGCCTTAGGAGATACCGATCTTTTGGGCAGA 985
Db 723 CACCTACGAGCCACCAAGGCGATAAAGTGCCTTAGGAGATACCGATCTTTTGGGCAGA 782
Qy 986 AGTAGAATCATGACTATACCACTATGGCGAAGAACTTAAATTTGGCGGGTAAAAACTAT 1045
Db 783 AGTAGAATCATGACTATACCACTATGGCGAAGAGCTCAAAATTTGGCGGGTAAAAACTAT 842
Qy 1046 CCGTAGGAGTATGGTTCAGAGCAATAGCCCTGATGAAAAACCCCTAGATTTAGTCATCAC 1105
Db 843 CCGTAGGAGTATGGTTCAGAGCAATAGCCCTGATGAAAAACCCCTTAGATTTAGTCATCAC 902
Qy 1106 TAACGCGATGATTATCGACTACACCGGATTTTACAAAGCCGACATTTGGGATTTAAAAACGG 1165
Db 903 CAACGCGATGATTATCGACTACACCGGATTTTAAAGCCGACATTTGGTATTAAAAATGG 962
Qy 1166 CAAAATCCATGGCAATTGGCAAGGCGAGGAAACAAAGGACATGCAAGATGGCGTAAGCCCTCA 1225
Db 963 CAAAATCCATGGTATTGGCAAGGCGGGAACAAAGACATGCAAGATGGCGTAAGCCCTCA 1022
Qy 1226 TATGCTCGTGGTGTGGGCACAGAAAGCACTAGCAGGGAAGTATGATTTATCCGCTGG 1285
Db 1023 TATGCTCGTGGTGTGGGCACAGAAAGCACTAGCAGGGAAGTATGATTTATCCGCTGG 1082
Qy 1286 GGGATCGATTACACACCCACTTCTCTTCTCCCAAAATTTCCCTACCGCTCTAGCCAA 1345
Db 1083 GGGATCGATTTCGACACCCACTTCTCTCTCCCAAAATTTCCCTACCGCTCTAGCCAA 1142
Qy 1346 TGGGTTTCAACCACTGTTTGGAGCGGCAAGGTCTCTAGATGGCAAGATGGCAACTAC 1405
Db 1143 TGGTGTTCACACCACTGTTTGGAGTGGCAAGGTCCGGTAGATGGCAAGATGGCACTAC 1202
Qy 1406 TATCACTCCGGGCAAAATGGAACTTGCAACCGCATGTTTGGCGCAGCAGAAAGATTTCTAT 1465
Db 1203 CATCACTCCGGGCAAAATGGAACTTGCAACCGCATGTTTGGCGCAGCAGTGAAGATTTCTAT 1262
Qy 1466 GAATGTGGCTTTTGGGCAAGCAATAGCTCTAGCAAAACAACTTGTAGAACAACT 1525
Db 1263 GAATGTGGCTTTTGGGCAAGCAATAGTCTTAGCAAAACAACTTGTAGAACAACT 1322
Qy 1526 AGAAGCGGCGCATTTGGTTTTAAATTCATGAAGACTGGGGCAACCAACCAAGTCGAT 1585
```

Db 1323 AGAAGCGGCGGATTTGGCTTTAAATTGATGAAGACTGGGCGACAAACCAAGTCCGAT 1382
QY CGATCACTGCTTTGAGCGTGGCAGATGAATACGATGTCGAAGTTGTATCCACACCGATAC 1645
Db CGATCACTGCTTTGAGCGTGGCAGATGAATACGATGTCGAAGTTGTATCCACACCGATAC 1442
QY AGTCAATAGGCGAGGTATGTAGATGACACCTAAATGAATGAACCGGCGGCGCATCCA 1705
Db GGTCAATAGGCGAGGTATGTAGATGACACCTAAATGAATGAACCGGCGGCGCATCCA 1502
QY TGCCTACCACTTTGAGGAGCGGTGGAGACACTCACCTGATGTTATCACCATGCGAGG 1765
Db TGCCTACCACTTTGAGGAGCGGTGGAGACACTCACCTGATGTTATCACCATGCGAGG 1562
QY CGAGCTCAATATTTCTCAAGCCGATCCACACCCCACTATTTCCTATACCAATTAATACGGT 1825
Db CGAGCTCAATATTTCTCAAGCCGATCCACACCCCACTATTTCCTATACCAATTAATACGGT 1622
QY TGCAGAACACTTAGACATGCTATGACATGCGACACCTAGACAAACGGATCCGCGAGGA 1885
Db TGCAGAACACTTAGACATGCTATGACATGCGACACCTAGACAAACGGATCCGCGAGGA 1682
QY TTTTACAATTTTCTCAAGCCGATCCACACCCCACTATTTCCTATACCAATTAATACGGT 1945
Db TTTTACAATTTTCTCAAGCCGATCCACACCCCACTATTTCCTATACCAATTAATACGGT 1742
QY TGATATGGGTGATCGGATGACAACTCGGATTCGCAAGCAATGGGCGCTGACGCGGA 2005
Db TGATATGGGTGATCGGATGACAACTCGGATTCGCAAGCAATGGGCGCTGACGCGGA 1802
QY AGTGATTCCTGAACTTTGCGAGCTCGGATGAAGATTAAGAAATTTGGTAAAGCTTCC 2065
Db AGTGATTCCTGAACTTTGCGAACTCGAGATGAAGATTAAGAAATTTGGTAAAGCTTCC 1862
QY TGAAGATGCAAGATGAAGATTTCCGATTAAGCGCTACATCTCCAATACACTAT 2125
Db TGAAGATGCGAGATGAAGATTTCCGATTAAGCGCTACATCTCCAATACACTAT 1922
QY CAACCCGCTTTGACCCAGCGGTGAGGATATATCGGCTCTGTCGAAGAGGCGCAAGAT 2185
Db TAAATCCGCTTTGACCCAGCGGTGAGGATATATCGGCTCTGTCGAAGAGGCGCAAGAT 1982
QY CGCGGACTTTGGTGTGGAACTCTGCTTTTGGCGTGAACCTAAGATTGTGATTA 2245
Db CGCGGACTTTGGTGTGGAACTCTGCTTTTGGCGTGAACCTAAGATTGTGATTA 2042
QY AGCGGTATGTTGCTTCTGAAATGGGCGATTTCAACGCTCTGTGCGCCACTCCCA 2305
Db AGGTGCGATGTTGCTTCTGAAATGGGCGATTTCAACGCTCTGTGCGCCACTCCCA 2102
QY ACCGCTTTATACCGGAAATGTTGGGATCAACGCGCAAGGCGAAATTTGACACCGCAT 2365
Db GCGGTTTATACCGGAAATGTTGGGATCAACGCGCAAGGCGAAATTTGACACCGCAT 2162
QY CACTTTTGTTCGATCGC 2385
Db CACTTTTGTGCTCAGCG 2182

RESULT 6

US-10-639-273-1
; Sequence 1, Application US/10639273
; Publication No. US20040142343A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; APPLICANT: Simpson, Kenneth W
; APPLICANT: Zhu, Jiaqian
; TITLE OF INVENTION: HELICOBACTER BIZZOERONII UREASE GENES AND THEIR USES
; TITLE OF INVENTION: IN DIAGNOSTIC AND TREATMENT METHODS
; FILE REFERENCE: 19603/3681
; CURRENT APPLICATION NUMBER: US/10/639, 273
; CURRENT FILING DATE: 2003-08-12
; PRIOR APPLICATION NUMBER: 60/404,337

; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 8407
; TYPE: DNA
; ORGANISM: Helicobacter bizzoeronii
US-10-639-273-1

Query Match 41.1%; Score 1184.4; DB 7; Length 8407;
Best Local Similarity 68.5%; Pred. No. 7.3e-260;
Matches 1681; Conservative 0; Mismatches 741; Indels 32; Gaps 2;

QY 183 AAATTAAACAAGGAGTAAAGTGAACCTCACACCCCAAGAGCAAGAAAGTCTTGTGT 242
Db AAATTGGTAGAAGGAGTTAGGATGAAATTAACCCCTAAGAGCTGCAAGCTCATGT 2378
QY 243 TATATATGCGGCGAAGTGGCTAGAAAGCAAGCAGAGGGCTTAAAGCTCAACCAAC 302
Db TGCATTTATGCGGCGAATTTGGCTTAAAAAACGCAAGCAAAATGGCGTTAAAGCTAAATATA 2438
QY 303 CCGAAGCCTATTGCTTACATTAGTGCCTATATTGGAAGAGCGCGCTGGAAAAA 362
Db CTGAGGAGTAGCCCTCATCAGTGCCTATGTGTAAGAGAGCCCGTGCAGGTAAAAA 2498
QY 363 CCCTTGGCCAGCTTATGGAAGAGTGCAATGCACATTTTTGAAAAAAGATCAAGTAATGCCCG 422
Db GTGTGGCGGATTTGATGCAAGAGGAGGACACATTTCTTAAAGCTGATGATGTCATGCCCG 2558
QY 423 GGTGGGTAAATATGTTTCCCGATCTAGGTGTAGAGCCACCTTTCTGTAGTGGTAGAAAC 482
Db GTGTAGCCCATATGATCCAGAAATGGGGATTTGAAGCTAACTTCCCTGATGGGCAAAAC 2618
QY 483 TTGTAAGTGAATTTGGCCCATCGAACAGATGAGCACTTCAAGCGGCGGAAGTAAT 542
Db TGTAAACCATTCATACCCCGTTGAAGATGGTGGGCATAAATTGGCTCCGGTGAAGTGA 2678
QY 543 TTGGTTGCGATAAAGACATCGAGCTCAATCAGGCAAGAAAGTAAACCGAACTTGAGGTGA 602
Db TTTTGAAGAACGAAGACATCACTTTGAATGAGGCAAGCAAGCCACCACTTTAGAGTGC 2738
QY 603 CTAATGAAGGCGCTAAATCTTGTGATGTTGGTAGCCATTTCCACTTTCTTGAAGTAAACA 662
Db ATAAACAAAGCGATCGCCCGTGCAGTGGGCTCCCACTTCCACTTCTTGAAGTGAATA 2798
QY 663 AGCAGCTAAATTCGATCGTGAAGAAAGCCTATGCAAAACGCTAGATATTTCCCTCTGCA 722
Db AGCTTTTGAATTTGATCGTGAAGAAAGCCTATGCAAAACGCTAGACATTTCTTCTGGAA 2858
QY 723 ACACGCTACGCAATTTGGGCGAGGACAAACCCGCAAGTGCAGTTGATTTCTTGTGGTGA 782
Db CCCTGTGCGCTTTGAACCGGTGAGAAAAAACCCGTGAATTTGATTCAAATTTGGCGTGA 2918
QY 783 GTAAAAAGTGTATGGCATGAACGGGCTTTGTAATAACATCGCGGATGAACGCCATAAAC 842
Db ACCAAGCATTTAGCGGCTTTAACTCTCTTGTGATCGCAAGCGGATCTGATGSCAAA 2978
QY 843 ATAAAGGCTTTGACAGGCGAAATCTCAGGATTTT----- 877
Db AACTTGTCTCTCAACCGCGCAAGAAACATGGCTTTTGTGTTTGTGAATTTGCGGTTGCGGATA 3038
QY 878 -----ATCAAGTAAAGGAGCTCCCATGAAATGAAAAA---ACAAGATATCTAAATACCT 930
Db AAAAAATAGGAAAGGACAAATCCGATGAAAAAAATCTCTCGAAAAAGAAATATGTTTCTATGT 3098
QY 931 ACGGACCCCAACAAAGGCGATAAAGTGGCTTTAGGAGATACCGATCTTTGGGCGAGAGTAG 990
Db ATGACCCCACTACGGGCGATTAAGTGAATTTGGCGGATACCGACTGTGATCTTAGAGTGC 3158
QY 991 AACATGACTATAACCTATGCGGAGAACTTAATAATTTGGCGGCTAAAACTATCCGTG 1050
Db AACATGACTGACCACTTATGGCGGAGAAATTAAGTTTGGTGGCGGTAACCACTTCGCG 3218

Db 187 AGCCAGAACTCGATCTTGTGCTCACTAAAGCCCTGATCGTGGATTACACCGGCATTTAT 246
Qy 1140 AAAGCCGATTTGGGATTTAAAAACGCAAAATCCATGGCATTTGGCAAGGACGAGAAACAAG 1199
Db 247 AAAGCCGATTTGGCATTTAAATAATGGCAAAATCCATGGCATTTGGCAAGGACGCAATAAA 306
Qy 1200 GACATGCAAGATGGCGTGAAGCCCTCATATGGTCTGGTGGTGGGACAGAACACTAGCA 1259
Db 307 GACATGCAAGATGGCGTGGTGAAGCAATCTTTGGTGGGCCCTGCTACTGAGGCTTTGGCC 366
Qy 1260 GGGGAAGGTATGATTTATACCCCTGGGGAATCGATTACACACCCACTTCTTCTCTCA 1319
Db 367 GCTGAAGGGCTGATTTGATACAGTGGTGGATGACACCCATCCACTTTATTTCTCCC 426
Qy 1320 CAACAATTTCCCTACCGCTCTAGCCATGGCGTTACAACCATGTTTGGAGCGGCACAGGT 1379
Db 427 CAACAATTTCCACAGCATTTGCCAGCGGATCAACAATGATTTGGTGGGGACAGGT 486
Qy 1380 CTTGTAGATGGACGAATGCGACTACTATCATCTCGGGCAAAATGGAATTTGCACCGCATG 1439
Db 487 CCAGCTGATGGGACTAAACGCACTACCATCACTCCGGGGCGCTGGAACCTTAAACCAATG 546
Qy 1440 TTGCGCGCAGCAGAGAGATTTCTATGATGTGGCTTTTGGGCAAAAGCAATAGCTCT 1499
Db 547 CTCGCTGCTCTGAAGAAATATGCCATGAATTTGGGCTATTTGGGTAAAGGGAATGTGCT 606
Qy 1500 AGCAAAAAACAATTTGTAGAAACAAGTAGAAGCGGCGCATTTGGTTTTAAATTTGCATGAA 1559
Db 607 TATGAACCTCTCCTGCTCGATCACTCGAAGCTGGAGCCATTTGGCTTTAAATCCACGAA 666
Qy 1560 GACTGGGCGACAACACCAAGTCGATCGATCACTGCTTGAGCGTGGCAGATGAATACGAT 1619
Db 667 GACTGGGCTAGCACACCTCGAGCCATCTACCATTTGCTTGAATGTGGCTGACAAATACGAT 726
Qy 1620 GTGCAAGTTTGTATCCACACCGATCAGATCAATGAGGAGGTATGTAGATGACACCTTA 1679
Db 727 GTGCAAGTGGCTATCCACACCGATCACTTGAATGAAGCGGGCTGTGTGGAAGACACTTTG 786
Qy 1680 AATGCAATGAACGGCGGCCATCCATGCTTACCACATTTAGGAGCGGGTGGAGGACAC 1739
Db 787 CAAGCCATTTGCTGGGCGCATTTCCACATTTTCCACACTGAAGGTGCTGCTGGCGGGCAC 846
Qy 1740 TCACCTGATTTATCACCATGGCAGCGAGCTCAATATTCTACCTCTCTCCACCAACCCCC 1799
Db 847 GCTCCGGATGTCATTAAGATGCTGGCGAATTAACATCTCCAGCTTCTACCAACCCC 906
Qy 1800 ACTATTCCTATACCATTAATACGTTGCGAGAACACTTAGACATGCTCATGACATGCCAC 1859
Db 907 ACCATTCTCTTCCCGTGAATACAGAAGCCGAACACATGGACATGTTGATGGTGGCCAC 966
Qy 1860 CACCTAGACAAACGGCATCCGCGAGGATTTACAATTTTCTCAAAGCGTATCCGCCCGGC 1919
Db 967 CACTTGGATTAATAACATCAAGAAGATGTCACAGTTTGTGATTTAGGATTCGCCCCCAA 1026
Qy 1920 TCTATCGCGCTGAAGATGCTCTCATGATATGGGTGTGATCGCGATGACAAAGCTCGGAT 1979
Db 1027 ACCATTCGCGCTGAGGACAAATCTCACGATATGGGGATTTCTCTATCACAGCTCTGAC 1086
Qy 1980 TCGAAGCAATGGGCGTGCAGCGGAAGTGAATTCCTCGAACTTTGGCAGACTGCGGATAAG 2039
Db 1087 TCCCAAGCGATGGGCGGTGAGCGAGGTCACTACCCGCACTTTGGCAAAACAGCGGACAAA 1146
Qy 2040 AATAAAGAATTTGGTAAAGCTTCTGAAATGGCAAAAGATTAACATTAATTTCCGCAT 2099
Db 1147 AACAAAAAGAAATTTGGTCTGCTTGGGAAAAAGGGGATAATGACAACTTCCGCATC 1206
Qy 2100 AAGCGCTACATCTCCAAATACACTATCAACCCGCTTTGACCCACCGCTGAGCGAGTAT 2159
Db 1207 AAGCGCTACATTTCCAAATACACCATCAACCCCGCTATTGACACCGGCAATTTCTGAAT 1266
Qy 2160 ATCGGCTCTGTGGAAGAGGCGAAGATCGCGCACTTGGTGGTGGGAATCCTGCTTTTTT 2219
Db 1267 GTCGCTCTGTAGAGTGGGCAAAATTCGCGGATTTGGTGTGTTGGAGTCTCTGCTTCTTT 1326

Qy 2220 GCGGTAAAAACCCAAATTCGTGATCAAAGGCGGTATGCTGCTTTCTCTGAAATGGCGAT 2279
Db 1327 GGCATTTAAACCCCAACATGATCATCAAAGGCGGATTCATCGCACTTTCTCAAATGGCGAT 1386
Qy 2280 TCTAACGCGTCTGTGCCACTCCCAACCGGTTTATTTACCGCGAAATGTTTGGGCAATCAC 2339
Db 1387 GCGAATGCGCTTATCCCCACTCCCAACCGGTTTATTTACCGCGAAATGTTTGGGCCCAT 1446
Qy 2340 GCGAAGCGGAAATTTGACACCGCATCACTTTGTGTTTCCAAAGTCGCTATGAAAAATGCG 2399
Db 1447 GGTAAAGCCAAATTTGACCAATATCATTTTGATATCCCAAGTGGCTTATGACACGGC 1506
Qy 2400 GTGAAAGAAAAGCTGGGCTTTAGAGCGCAAGTTTCTACCGGTCAAAAACTGCCGTAACATC 2459
Db 1507 ATTTAAAGAGAGTTGGGCTTGGCAAGAGTGTTTGTTCAGTTTAAATACTGCGCAACATC 1566
Qy 2460 ACCAAGAAAGACTTCAAGTTCAAGCAAAAACGCAAAATACCGTCGATCCGAAACACC 2519
Db 1567 ACCAAAAAGACCTCAAAATTCACAGATTTTACCGCACACATCGAAGTCAATCTCTGAAACC 1626
Qy 2520 TTGAGGCTCTTTGTAGATGCAAACTCTGCACCTCTTAAACCCACTCGCAAGTGCCTCTA 2579
Db 1627 TACAAAGTTAAAGTGGATGGCAAGAGGTACTTTCCAAAGCAGCGATATAAATCAGCCTA 1686
Qy 2580 GCCAGCGCTACACTTTTCTTTCTAG 2603
Db 1687 GCACAACTCTACAACTTTGTTCTAG 1710

RESULT 8
US-10-476-313-10
; Sequence 10, Application US/10476313
; Publication No. US20040241175A1
; GENERAL INFORMATION:
; APPLICANT: MCKENZIE, BRENT
; APPLICANT: BOYLE, JEFFEREY
; APPLICANT: LEW, ANDREW
; TITLE OF INVENTION: Antigen Targeting
; FILE REFERENCE: BDWP-002
; CURRENT APPLICATION NUMBER: US/10/476,313
; CURRENT FILING DATE: 2003-10-24
; PRIOR APPLICATION NUMBER: PR5241
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 1719
; TYPE: DNA
; ORGANISM: Helicobacter felis
US-10-476-313-10

Query Match 31.3%; Score 902.4; DB 8; Length 1719;
Best Local Similarity 70.6%; Pred. No. 1.1e-195;
Matches 1203; Conservative 0; Mismatches 501; Indels 0; Gaps 0;

Qy 900 AAATGAAAAACAGNATATGTAAATACCTACCGACCCCAAGCGATAAAGTGGCG 959
Db 15 AAGATTTCAAGAAAGAAATATGTTTCTATGTTGTTCCACTACCGGGGATCGTGTAGA 74
Qy 960 TTAGGAGATACCGATCTTTGGGCAAGATAGAACATGACTATACCACTATGGCGAAGAA 1019
Db 75 CTCGGGACACTGATTTGATCTTAGAAGTGGAGCATGATTCACCATTTATGGTGAAGAG 134
Qy 1020 CTTTAAATTTGGCGGGTAAAACTATCCGTGAGGGTATGGGTGAGAGCAATAGCCCTGAT 1079
Db 135 ATCAAATTTGGGGCGGTAAAACTATCCGTGATGGATGAGTCAAAACCAATAGCCCTAGC 194
Qy 1080 GAAAAACACCTAGATTTAGTTCATCATCTAAACGCGATGATTTATCGACTACACCGGATTTAC 1139
Db 195 TCTTATGAATTTAGATTTGGTGTCTCACTAACGCCCTCATTTGTGGACTATACGGGCATTTAC 254
Qy 1140 AAAGCCGACATTGGGATTTAAAAACGGCAAAATCCATGCGATTTGGCAAGCAGGAAACAAG 1199

Db 255 |||||AAAGCCGACATTGGGATTAAGACGCGAAGATTGCGAGCATTTGGCAAGGACGGCAATAAG 314
Qy 1200 GACATCAAGATGGCGCTAAGCCCTCATATGGTGGTGGTGGGACAGAGAGCACTAGCA 1259
Db 315 GACATCAAGATGGCGCTAGATATATCTTTGGTAGTCTCTGCTACAGAGGCTTTGGCA 374
Qy 1260 GGGAAAGGTATGATTAATACCGCTGGGGAAATCGATTCAACACCCATCTTCCTTTCCCA 1319
Db 375 GCTGAGGGCTTGATTGTAACCGCTGGTGCGATCGATACGCATATTCACTTTATCTCTCC 434
Qy 1320 CAACAATTCCTTACCGCTTAGCCATGGCTTACACCAATGTTTGGAGCGGCACAGGT 1379
Db 435 CAACAATTCCTTACTGCTTTGCGAGCGGGTTACCAACCATGATTGGAGGAGGCACAGGA 494
Qy 1380 CTTGTAGATGGCAGCAATCGGACTACTCACTCCGGGCAATGGAATTTGCCCGCATG 1439
Db 495 CTTGCGGATGGCAGCAATGGACCAATCACTCCGGACGCGTAACTCTAAAGATG 554
Qy 1440 TTGCGCGCAGCAGAGATATTCTATGAATGTGGGCTTTTGGGCAAGGCAATAGCTCT 1499
Db 555 TTGCGTGCAGCGAAGATACGCATGAACTAGGCTTTTGGCTAAGGGGAATGTGCT 614
Qy 1500 AGCAAAAAACAATCTGTAGAACAAAGTAGAAGCGGGCGCATTTGTTTAAATTTGCATGAA 1559
Db 615 TACGAACCTCTTTTACGCGATCAGATTGAAGCAGGGGCGATTGGTTTTTAAATCCACGAA 674
Qy 1560 GACTGGGCGACACACCAAGTCGATCGATCTGCTTGGAGCTGGCGAGATGAATACGAT 1619
Db 675 GACTGGGGAAGCACACCTCGAGCTATTCACCATGCTCAATGTGCGCGATGAATACGAT 734
Qy 1620 GTGCAAGTTTGTATCCACACCGCATACAGTCAATGAGGCGAGTTATGTAGATGACACCTTA 1679
Db 735 GTGCAAGTTGCTATCCACACCGATACCTTAAAGGCGGCTGTGTAGAAGACACCTTA 794
Qy 1680 AATGCAATGAACGGGCGCGCATTCATGCTTACCATTCAGGAGCGGTTGAGGACAC 1739
Db 795 GAGGCGATTTCGCGGCGCACCATCCATACCTCCACACTGAAGGGCTGGGGTGGACAC 854
Qy 1740 TCACCTGATGTTATCACCATTGGCAGGAGCTCAATATTCTACCTCTCCACACCCCC 1799
Db 855 GCTCCAGATGTTATCAAAATGGCAGGGGAATTTAACATTTCTACCCGCTCTACTAACCCG 914
Qy 1800 ACTATTCCTATACCATTAATACGTTGCAAGACACTTACACATGCTCATGACATGCCAC 1859
Db 915 ACAATTCCTTTACCAAAAAACACTGAAGCCGAGCACATGACATGTTAATGGTGGCCAC 974
Qy 1860 CACCTAGACAAACGCGATCCGCGAGGATTTACAATTTTCTCAAGCCGTTATCCGCCCGGC 1919
Db 975 CACTTGGATTAAGTATCAAGGAAGATGTCAGTTTCCCGATTTCGAGGATTCGCCCCAA 1034
Qy 1920 TCTATCCGGCTGAAGATGTGCTCCATGATATGGGTGTGATTCGGATGACAAGCTCGGAT 1979
Db 1035 ACTATCCGGCTGAAGACCAATCCATGACATGGGATCTTTTCTATCACCAGCTCCGAC 1094
Qy 1980 TCGAAGCAATGGGGCTGCAGCGGAAGTATCTCTCGAATTTGGCAGACTGCGGATAAG 2039
Db 1095 TCTCAGGCTATGGAGCGCGTAGCGAGGTGATCACGCGACTTTGGCAGACGAGCAAAA 1154
Qy 2040 AATAAAAAAGAAATTTGTTAAGTTCCTGGAAGATGGCAAGATTAACGATAATTTCCGCATT 2099
Db 1155 AACAAAAAGATTTGGGCGCTTGAAGAGAAAAAGGCGATTAACGAACTTTCCGCATC 1214
Qy 2100 AAGCGCTACATCTCAAATACACTATCAACCCCGCTTTGACCCACGCGCTGAGCGAGTAT 2159
Db 1215 AAACGCTACATCTTAATACACCATCAACCCCGGATCGCGATGGGATTTCTGACTAT 1274
Qy 2160 ATCGGCTCTGTGAAGAGGCAAGATCGCCGACTTGGTGTGTGGAATCTCGCTTTTTT 2219
Db 1275 GTGGGCTCTGTGAAGTGGGCAATACGCCGACCTCGTGTCTTTGGAGTCCGGCTTTCTTT 1334
Qy 2220 GCGTAAACCCCAAAATCGTGATCAAGGCGGTATGGTGTCTTCTCTGAATGGCGGAT 2279
|||

Db 1335 GGCATTAAAGCCCAATATGATTTAAGGGCGGATTTATTTCGCTCTCTCAAATGGCGAT 1394
Qy 2280 TCTAACCGCTGTGGCCCACTCCCAACACCGGTTTATTATCCGCGAAATGTTTGGGCAATCAC 2339
Db 1395 GCCAATCGCTATTCTCCACCCCTCAGCCCGCTATTATACCGGTGAAATGTTTGGACACCAT 1454
Qy 2340 GGCAGGCGAAATTTGACACCAGCATCACTTTGTTTCCAAAGTCGCTTATGAAATGCG 2399
Db 1455 GGGAAAAACAATTCGACACCAATATCACTTTCGTGTCCCAAGCGGCTTACAAGGCGGG 1514
Qy 2400 GTGAAAGAAAAAGCTGGGCTTAGAGCGCCAAAGTTCTTACCGGTCAAAAACTGCCGTAAACATC 2459
Db 1515 ATCAAGAGAACTAGGGCTAGATCGGTGTATTGCCAGTGAAGAACTGTGCAATATC 1574
Qy 2460 ACCAAGAAAGACTTTCAAGTTCAACGACAAAAACGGCAAAAAATCACCGTGCATCCGAAAAAC 2519
Db 1575 ACTAAAAAGGACCTCAAAATTCACGATGTGACCGCACATATTGATGTCAACCTGAAAC 1634
Qy 2520 TTCGAGTCTTTGTAGATGGCAAACTGTGACCTCTTAAACCCACCTCGCAAGTGCCTCTA 2579
Db 1635 TATAGGTGAAAGTGGATGGCAAGAGGTAAACCTCTAAAGCAGCAGATGAATTGAGCCTA 1694
Qy 2580 GCCAGCGCTACACTTTCTTCTAG 2603
Db 1695 GCGCAACTTTATAATTTGTTCTAG 1718

RESULT 9

US-10-335-977-3849
Sequence 3849, Application US/10335977
Publication No. US20040052799A1

GENERAL INFORMATION:

APPLICANT: DOUGLAS SMITH et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
RELATING TO HELICOBACTER PYLORI FOR
DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 10031

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD

STREET: 28 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: CD/ROM ISO9660

OPERATING SYSTEM: Windows NT 4.0

SOFTWARE: UNIX

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/335,977

FILING DATE: 30-Dec-2002

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/993,002

FILING DATE: 17-DEC-1997

ATTORNEY/AGENT INFORMATION:

NAME: Mandragouras, Amy E.

REGISTRATION NUMBER: 36,207

REFERENCE/DOCKET NUMBER: GTN-018

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400

TELEFAX: (617)742-4214

INFORMATION FOR SEQ ID NO: 3849:

SEQUENCE CHARACTERISTICS:

LENGTH: 1710 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: circular

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Helicobacter pylori

| | | | | | | | | | |
|---|------|---|------|--|--|--|--|--|--|
| FEATURE: | | | | | | | | | |
| NAME/KEY: misc feature | | | | | | | | | |
| LOCATION: (B) LOCATION 1...1710 | | | | | | | | | |
| SEQUENCE DESCRIPTION: SEQ ID NO: 3849; | | | | | | | | | |
| US-10-335-977-3849 | | | | | | | | | |
| Query Match | | | | | | | | | |
| Best Local Similarity 30.7%; Score 886.4; DB 7; Length 1710; | | | | | | | | | |
| Matches 1193; Conservative 0; Mismatches 511; Indels 0; Gaps 0; | | | | | | | | | |
| QY | 900 | AAAAAGMAAAACAGAAATATGTAATACCTAGGACCCACCAAAAGCGGATAAAGTCGC | 959 | | | | | | |
| DB | 7 | AAGATTAGCAAAAAGAAATATGTTCTATGTATGGCCCTACCACAGCGGAGTAAGTGAGA | 66 | | | | | | |
| QY | 960 | TTAGGAGATPACCGATCTTTTGGGAGAGTAGAACATGACTATACCACTATGGCGGAAGAA | 1019 | | | | | | |
| DB | 67 | TTGGCGGATACAGATTGTATGCTGAGTAGAACATGACTACACCATTTATGGCGAAGAG | 126 | | | | | | |
| QY | 1020 | CTTAAATTTGGCGCGGTAAACTATCCGTGAGGGTATGGGTGAGAGCAATAGCCCTGAT | 1079 | | | | | | |
| DB | 127 | CTTAAATTCGGTGGCGGTAAACCCCTAAGAGAGGCGATGAGCCAATCCCAACACCCCTAGC | 186 | | | | | | |
| QY | 1080 | GAAPACACCTAGATTAGTCACTACCTAAGCGGATGATTATCGACTACACCGGGATTAC | 1139 | | | | | | |
| DB | 187 | AAAGAGAACTGGATTGTATCATCACTAAGCTTTAATCGTGGAATTACACCGGTATTTAT | 246 | | | | | | |
| QY | 1140 | AAAGCGGACATTTGGGATTAAAAACGGCAAAATCCATGGCATTTGGCAAGGAGGAAACAAG | 1199 | | | | | | |
| DB | 247 | AAAGCGGATTTGGTATTAAAGCGGCAAAATCGCTGGCATTTGGTAAAGCGGTAAACAA | 306 | | | | | | |
| QY | 1200 | GACATGCAAGATGGCGTAAAGCCCTCATATGGTCGTGGGTGTGGGCAACAGAACCTAGCA | 1259 | | | | | | |
| DB | 307 | GACATGCAAGATGGCGTTAAAAACAATCTTAGCGTGGTCCCTACTAGGAGCGGTAGCT | 366 | | | | | | |
| QY | 1260 | GGGGAAGGTATGATTATTACCGCTGGGGAAATCGATTACACACCCACTTCCTTTCCCA | 1319 | | | | | | |
| DB | 367 | GGTGAAGGCTTAATCGTAAACGGCTGGTGTATTGACACACACATCCACTTTCAITTCACCC | 426 | | | | | | |
| QY | 1320 | CAACAATTCCTTACCGCTCTAGGCAATGGCTTACCAACCATGTTTGGAGCGGCACAGGT | 1379 | | | | | | |
| DB | 427 | CAACAATTCCTTACAGCTTTTGCAGCGGTGTAAACCAACATGATTGGTGGCGGAACCGGT | 486 | | | | | | |
| QY | 1380 | CCTGTAGATGGCAATGCGACTACTATCACTCCGGGCAAAATGGAACTTGCACCGCATG | 1439 | | | | | | |
| DB | 487 | CCTGTCTGTATGGCACTAAATGCGACTACTATCACTCCAGGTAGAGAACAATTTAAATGGATG | 546 | | | | | | |
| QY | 1440 | TTGGCGGAGCAGAGAGATTTCTTATGAATGTGGGCTTTTGGGCAAAAGGCAATAGCTCT | 1499 | | | | | | |
| DB | 547 | CTCAGAGCGGCTGAAGAAATATCTATGAACCTTAGGTTTCTTAGCTAAAGGTAACGGTTCT | 606 | | | | | | |
| QY | 1500 | AGCAAAAACAACCTTGTAGAACAGTAGAGCGGCGGATTTGGTTTAAATTTGCATGAA | 1559 | | | | | | |
| DB | 607 | AACGACGCGAGCTTAGCGGATCAAAATGAAGCGCGGTGCGATTGGCTTTAAATTTCCACGAA | 666 | | | | | | |
| QY | 1560 | GACTGGGGCACAAACCAAGTCGATCGACTCTTGAGCGTGGCAGATGAATACGAT | 1619 | | | | | | |
| DB | 667 | GACTGGGGCACCACTCTTCTGCAATCAATCATGCGTTAGACGTTGGCGCAATACGAT | 726 | | | | | | |
| QY | 1620 | GTGCAAGTTTGTATCCACACCGGATACAGTCAATGAGCGAGGTTATGTPAGATGACACCTTA | 1679 | | | | | | |
| DB | 727 | GTGCAAGTCGCTATCCACACAGACACTTTGAATGAAGCCGGTTGCGTGGGAAGACACTATG | 786 | | | | | | |
| QY | 1680 | AATGCAATGAACGGGGCGCCATCCATGCCATACCAATTGAGGAGCGGTGGAGGACAC | 1739 | | | | | | |
| DB | 787 | GCAGCTATTGGCGGAGCGCATATGCACTTTCCACACTGAAGCGCGTGGTGGCGGACAC | 846 | | | | | | |
| QY | 1740 | TCACTGTATGTTATCACATGGGAGCGGAGCTCAATATTCTACCTCTCCACACACCCCC | 1799 | | | | | | |
| DB | 847 | GCTCTGACATATTAAAGTGGCGGTGAACACAACTCTACCCGTTTCACTAACCCC | 906 | | | | | | |
| QY | 1800 | ACTATTCCCTATACCAATTAATACGGTTGCAAGACACTTACATGATGCTCATGACATGCCAC | 1859 | | | | | | |
| DB | 907 | ACTATCCCTTTCACTGTGAATACAGAAGCGGAAACACATGACATGCTCATGGTGTGCCAC | 966 | | | | | | |

| | | | | | | | | | |
|----|------|--|------|--|--|--|--|--|--|
| QY | 1860 | CACCTAGACAAAACGCATCCCGAGGATTTACAAATTTTCTCAAAGCCGATCCGCCCCCGC | 1919 | | | | | | |
| DB | 967 | CACCTGGATAAAAGCAATTAAGAAGATGTCCAGTTCCGTGATTCAAGGATCCGCCCTCAA | 1026 | | | | | | |
| QY | 1920 | TCATCCGCGCTGAAGATGTCTCCATGATATGGTGTGATCGGATGACAAGCTCGAT | 1979 | | | | | | |
| DB | 1027 | ACCATTTCCGCGCTGAAGACACTTTTGCATGACATGGGATTTTCTCAATCACCAGTTCTGAC | 1086 | | | | | | |
| QY | 1980 | TCCCAAGCAATGGGCGTCAGGCGAAGTGATTCCTCGAACTTGGCAGACTGCGGATAG | 2039 | | | | | | |
| DB | 1087 | TCACAGCGATGGTGTGTGGTGAAGTTATCACTAGGACTTGGCAAAACAGCTGACAAA | 1146 | | | | | | |
| QY | 2040 | AATAAAAAAGAAATTTGGTAAAGCTTCTGAAAGATGGCAAGATAACGATAAATTTCCGCATT | 2099 | | | | | | |
| DB | 1147 | AACAAAAAGAAATTTGGCCCTTGAAAAGAAAGAAAGGCGATAACGACAACTTCAGGATC | 1206 | | | | | | |
| QY | 2100 | AAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACCCAGCGGTGAGCGAGTAT | 2159 | | | | | | |
| DB | 1207 | AAAACGCTACTTGTCTAAATACACCATTAACCCAGCGATCGCTCATGGGATTTAGCGAGTAT | 1266 | | | | | | |
| QY | 2160 | ATCGGCTCTGTGAAGAGGCAAGATCCGACATTTGGTGGTGTGGAATCCCTGCTTTTTT | 2219 | | | | | | |
| DB | 1267 | GTAGTTCTGTAGAAGTGGGCAAGTGCGTGACTTTGGTATTGTGGAGTCCAGCATTCITT | 1326 | | | | | | |
| QY | 2220 | GGCGTAAAAACCCAAATCGTGATCAAAAGGCGGTATGGTGTCTTCTCTGAAAATGGGCGAT | 2279 | | | | | | |
| DB | 1327 | GGCGTGAACCCCAACATGATCATCAAAAGGCGGATTCATTGCGTTAAGCCAAATGGGCGAT | 1386 | | | | | | |
| QY | 2280 | CTAAACGCGTCTGTGCCACTCCCAACCGGTTTATTACCCGCAAAATGTTTGGGCGATCAC | 2339 | | | | | | |
| DB | 1387 | GCBAACGCTTCTATCCCTACCCCAACACCGGTTTATTACAGAGAAATGTCGCTCACCAT | 1446 | | | | | | |
| QY | 2340 | GGCAAGCGGAAATTTGACACGACATCACTTTTCTTCCAAAGTCGCTATGAAATGCG | 2399 | | | | | | |
| DB | 1447 | GGTAAAGCTAAATACGATGCAACATCACTTTTGTGCTCAAGCGGCTTATGACAAAGC | 1506 | | | | | | |
| QY | 2400 | GTCAAGAAAAAGCTGGGCTTAGAGCGCAAGTCTTACCGGTCAAAAACTGCGCTAACATC | 2459 | | | | | | |
| DB | 1507 | ATTAAAGAGNAATTAGGCTTGAAGACAGTGTTCGCGTTAAAAATTCAGAAACATC | 1566 | | | | | | |
| QY | 2460 | ACCAAGAAAGACTTCAAGTTCAACAGCAAAAACGGCAAAAATCAACGTCGATCCGAAAACC | 2519 | | | | | | |
| DB | 1567 | ACTAAAAAGACATGCAATTCATGACACTACCGCTCACTTGAAGTCAATCCTGAAACT | 1626 | | | | | | |
| QY | 2520 | TTGAGGCTTTGTAGATGGCAAACTCTGCACCTCTAAACCCACTCCGNAAGTGCCTCTA | 2579 | | | | | | |
| DB | 1627 | TACCATGTGTGCTGGATGGCAAGAAAGTCACTTCTTAAACCAAGCTAATAAAGTGAGCTTG | 1686 | | | | | | |
| QY | 2580 | GCCGAGCGCTACACTTTTCTCTAG | 2603 | | | | | | |
| DB | 1687 | GCTCACTCTTTAGCATTTTCTAG | 1710 | | | | | | |

RESULT 10

US-09-895-913A-251
; Sequence 251, Application US/09895913A
; Patent No. US20020160456A1
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean Francois
; APPLICANT: Omen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding No. US20020160456A1el Helicobacter Polypeptides in the
; FILE REFERENCE: 06132/043002
; CURRENT APPLICATION NUMBER: US/09/895,913A
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 08/881,227
; PRIOR FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 368

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 251
; LENGTH: 1315
; TYPE: DNA
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (51)...(1757)
US-09-895-913A-251

```

| Query Match | 30.7%; | Score 885.8; | DB 3; | Length 1815; |
|-----------------------|-----------------|---|-----------|--------------|
| Best Local Similarity | 70.0%; | Pred. No. 7.2e-192; | | |
| Matches 1193; | Conservative 0; | Mismatches 512; | Indels 0; | Gaps 0; |
| QY | 900 | AAAATGAAAAAACAAGAAATATGTAAATACCTACGCGACCCACCAAGCGCATAAAGTCGCG | 959 | |
| DB | 57 | AAGATTAGCAGAAAAGAAATATGTTCTATGTATGCCCTACTACAGCGGATAAAGTGA | 116 | |
| QY | 960 | TTAGGAGATACCGATCTTTGGCGAGAAAGTAGAACATGACTATATCCACTATGGCGAAGAA | 1019 | |
| DB | 117 | TTGGCGGATACAGACTTGTATCGCTGAAGTAGAACATGACTACACCAATTTATGGCGAAG | 176 | |
| QY | 1020 | CTTAAATTTGGCGCGGTAAACATATCCGTGAGGATATGGTACAGCAATAGCCCTGAT | 1079 | |
| DB | 177 | CTTAAATTCGTGGCGGTAAACCCCTGAGAGAAGGCATGAGCCCAATCCAAACCCCTAGC | 236 | |
| QY | 1080 | GAATAACACCTAGATTTAGTCATCACGTAAACGCGATGATTATCGACTACACCGGATTTAC | 1139 | |
| DB | 237 | AAAGAAGATTGGATCTTAATCATCATACGCTTTAATCTGTGGATTACCGGTATTTAT | 296 | |
| QY | 1140 | AAAGCCGACATTTGGGATTTAAAAACGGCAAAATCCATGGCATTTGCAGAGCGAGAAACAAG | 1199 | |
| DB | 297 | AAAGCGGATTTGGTATTTAAAGATGGCAAAATCGCTGCATTTGGTAAAGCGGTAAACAA | 356 | |
| QY | 1200 | GACATGCAAGATGGCGTAAGCCCTCATATGGTCTGTGGGTGTGGCAACAGAACCATAGCA | 1259 | |
| DB | 357 | GACATGCAAGATGGCGTTAAAAACAATCTTTAGCGTAGGTCTCTGCTACTGAAGCCCTTAGCC | 416 | |
| QY | 1260 | GGGGAAGGTATGATTATTACCGCTGGGGGAATCGATTACACACCCCACTTCTTTCTCCA | 1319 | |
| DB | 417 | GGTGAAGGTTTGATCTGTAACCTGCTGGTGGTATTGACACACATCCACTTCATTTTCACCC | 476 | |
| QY | 1320 | CAACAATPTCCCTACCGCTCTAGCCAAATGGCGTTTACAACCATGTTTGAGGCGGCACAGGT | 1379 | |
| DB | 477 | CAACAATCCCTACAGCTTTTGAAGCGGTGTAAACACCATGATTGGTGGCGGAACCTGGT | 536 | |
| QY | 1380 | CTGTAGATGGCAGGAATGGCACTATCATCTCCGGGCAAAATGGAACTTTGCAACCGCATG | 1439 | |
| DB | 537 | CCTGCTGATGGCACTAATGGCACTACTATCACTCCAGCGAAGAAATTTAAAAATGGATG | 596 | |
| QY | 1440 | TTGCGGCGACGAGAAGTAGTCTCTCAATGTGGGCTTTTGGSCAAAGCAATAGCTCT | 1499 | |
| DB | 597 | CTCAGAGCGCTGAAGAATATTCTATGACCTTAGGTTCTTTGGCTAAAGGTATACGCTTCT | 656 | |
| QY | 1500 | AGCAAAAAACAACCTGTGTAGAACAGTAGAAGCGGCGCGATTTGGTTTTAAATTTGCATGAA | 1559 | |
| DB | 657 | AACGACGCGAGCTTAGCCGATCAAAATGAAGCTGGTGGATGGCTTTAAAAATCCACGAA | 716 | |
| QY | 1560 | GACTGGGGCAACAACAAGTGGCATCGATCATCTGTTTGAGCTGGGAGATGAATACGAT | 1619 | |
| DB | 717 | GACTGGGGCAACCACTCTTCTGCAATCAATCATGCGTTAGATGTTGTGAGCACAATATAGAT | 776 | |
| QY | 1620 | GTCCAGTTTGTATCCACACCGATACAGTCAATGACGCGAGGTTATGTAGATGACACCCCTA | 1679 | |
| DB | 777 | GTGCAAGTTCGCTATPCCAACAGACACTTTGAATGAAGCCGTTGCGTGGGAAGACACTATG | 836 | |
| QY | 1680 | AATGCAATGAACGGCGCGCCATCCATGCTCATCCACAATTGAGGAGCGGGTGGAGGACAC | 1739 | |
| DB | 837 | GCAGCTATTGCCGACGCCATATGCAACATTTCCACACTGAAGTGTCTGGCGGGACAC | 896 | |
| QY | 1740 | TCACCTGATGTTATCACCATGTGGACGGCGAGCTCAATATTTCTACCCCTCTCTCCACCCCCC | 1799 | |
| DB | 897 | GCTCCTGATATTTAAAGTAGTGGTGAACAACAACATTTCTCCGCTTCTCACTTAACCCCC | 956 | |

| | | | |
|----|------|---|------|
| QY | 1800 | ACTATTCCTTACATCACTTAATACGGTTGCGAAGACCACTTAGACATGCTCATGACATCGCCAC | 1859 |
| DB | 957 | ACTATCCCTTTCTCACTGTGAAATACAGAAGCAGAACACATGGACATGCTTATATGGTGTGCCAC | 1016 |
| QY | 1860 | CACCTAGACAAAACGATCCGCGAGGATTTTACAATTTTTCTCAAAAGCGTATCCCGCCCGCGC | 1919 |
| DB | 1017 | CACTTGGATTAAGCATTAAGAAGATGTTTCAGTTGCTGATTTCAAGGATCCGCGCTCAA | 1076 |
| QY | 1920 | TCTATCCGGCTGAAGATGTGCTCCATGATATGGGTGTGATCCGGATGACAAGCTCGGAT | 1979 |
| DB | 1077 | ACCATTCCGGCTGAAGACACTTTTGTCATGACATGGGGATTTTTCTCAATCACCAAGCTCTGAC | 1136 |
| QY | 1980 | TCGCAAGCAATGGGCGTGCAGCGAAGTCAATTCCTCGAAGCTTTGGCAGAGCTCGCGGATAAG | 2039 |
| DB | 1137 | TCTCAAGCTATGGGTGCTGTGGGTGAAGTTATCACTAGAACCTTTGGCAACAGCTGACAAA | 1196 |
| QY | 2040 | AATAAAAAAGAAATTTGGTAAAGCTTCCTGAAAGATGGCAAGATAACGATAATTTCCGCATT | 2099 |
| DB | 1197 | AACAAAAAGAAATTTGGCCGCTTGAAAGAAGAAAAGGCGATAAACGACAACTTCAGGATC | 1256 |
| QY | 2100 | AAGCGCTACATCTCCAAATACATATCAACCCCGCTTTTGACCAACGGCGTGAGCGAGTAT | 2159 |
| DB | 1257 | AAACGCTACTTGTCTAAATACACCATTAACCCAGCGATCGCTCATGGGATTAGCGAGTAT | 1316 |
| QY | 2160 | ATCGGCTCTGTGGAAGAGGCAAGATCGCCGACTTGGTGGTGTGGAATCCTGCTCTTTTTT | 2219 |
| DB | 1317 | GTAGGTTCTGTAGAAGTGGCAAGTGCGCTGACTTGGTATTTGTGGAGTCCCGCATCTTTT | 1376 |
| QY | 2220 | GGCGTAAACCCAAAATCGTGATCAAAAGCGGTATGGTGGTCTTCTCTGAATGGCGCAT | 2279 |
| DB | 1377 | GGCGTAAACCCCAACATGATCATCAAGGCGGCTTCATTGCGTTGAGTCAATGGCGTGAC | 1436 |
| QY | 2280 | TCTAAACGGTCTGTGCCACTCCCAACCGGTTTATTACCGCAAAATGTTTGGGATCAC | 2339 |
| DB | 1437 | GCGAACGCTTCTATCCCTACCCCAACACCAAGTTTATTACAGAGAAATGTTTCGCTCATCAT | 1496 |
| QY | 2340 | GGCAAGCGGAAATTTTGACACGAGCATCACTTTTTTTGTTTTTCAAAGTCGCTATGAAATGGC | 2399 |
| DB | 1497 | GGTAAAGCAANTACGATGCAACATCACTTTTGTGTCTCAAGCGCTTATGACAAGGC | 1556 |
| QY | 2400 | GTGAAGAAAGCTGGGCTTTAGAGCGCCAAAGTTCTACCGGTCAAAAATCTGCCGTAAACATC | 2459 |
| DB | 1557 | ATTAAGAAGAAATTAGGGCTTGAAGAACAAGTGTTCGCGGTAAAAAATTTGCAAAAACATC | 1616 |
| QY | 2460 | ACCAAGAAAGACTTCAAGTTTCAACGACAAACGGCAAAATACACCGTCGATCCGAAAACC | 2519 |
| DB | 1617 | ACTAAAAAAGACATGCAAAATTTCAACGACACTTACCCTGCACATTTGAAGTCAATCTCGAACT | 1676 |
| QY | 2520 | TTCCAGGTCTTTGTAGATGGCAAACTCTGCACCTCTAAACCCCACTCGCAAGTGCCTCTA | 2579 |
| DB | 1677 | TACCATGTGTTCTGGTATGGCAAGAAGTAACTTCTTAAACCGCAATAAAGTGAGCTTG | 1736 |
| QY | 2580 | GCCACGCGCTACACTTTCTTCTTAGG | 2604 |
| DB | 1737 | GCGCAACTCTTTAGCATTTTCTAGG | 1761 |

RESULT 11
US-10-282-122A-22427
; Sequence 22427, Application US/10282122A
; Publication No. US2004029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert

APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA 034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 22427
LENGTH: 1710
TYPE: DNA
ORGANISM: Helicobacter pylori
US-10-282-122A-22427

Query Match 30.7%; Score 884.8; DB 7; Length 1710;
Best Local Similarity 70.0%; Pred. No. 1.2e-191;
Matches 1192; Conservative 0; Mismatches 512; Indels 0; Gaps 0;

QY 900 AAAATGAAAACAGAAATATGTAATACCTACGACCCACCAAGGGCGATAAAGTGGC 959
DB 7 AAGATTAGCAGAAAAGAAATATGTTCTATGTATGGCCCTACTACAGGCGATAAAGTGAGA 66

QY 960 TTAGGAGATACCGATCTTTGGGCGAGAGTAGACATGACTATACCACTATGCGGAGAA 1019
DB 67 TTGGGCGATACAGACTTGATCGCTGAAGTAGAATGACTACACCATTTATGGCGAAGAG 126

QY 1020 CTTAAATTTGGCGCGGTAAACCTATCCGTAGGGGTATGGGTGAGAGCAATAGCCCTGAT 1079
DB 127 CTTAAATTCGGTGGCGGTAAACCTTGAGAGAGGCGATGAGCCAAATCCAAACCTTAGC 186

QY 1080 GAAAACACCTAGATTAGTATCATCTAACGCGATGATTATGACTATACACCGGATTTAC 1139
DB 187 AAAGAGAATTTGGATCTGATCATCACTAAACGCTTTAATCGTGATTTACACCGGTATTTAT 246

QY 1140 AAAGCCGATTTGGGATTTAAACGCGCAATCCATGGCATTTGGCAAGCGAGAAACAG 1199
DB 247 AAAGCGGATTTGGTATTAAGATGCAAAATCGCTGGCATTTGGTAAAGCGGTAAACAA 306

QY 1200 GACATGCAAGATGGCGGTAGCCCTCATATGGTGGTGGGTGGGCGACAGAACCTAGCA 1259
DB 307 GACATGCAAGATGGCGGTAAACAAATCTTAGGCTAGGTCTGCTACTGAAGCTTAGCC 366

QY 1260 GGGGAAGGTATGATTATACCGCTGGGGGAATCGATTACACACCCATCTCTCTTCTCCA 1319
DB 367 GGTGAAGGTTTGATCGTAACTGCTGGTGGTATTGACACACATCCACTTTCACTTTCACCC 426

QY 1320 CAACAATTCCTACCGCTTAGCCATGGGTTTAAACCAATGTTTGGAGCGGCGACAGGT 1379
DB 427 CAACAATTCCTACCGCTTTGGAAGCGGTGTAACCAATGATTGGTGGCGGAACGTGT 486

QY 1380 CCTGTAGTGGCAGGATGGACTACTATCACTCCGGGCAAAATGGAATTTGACCGCATG 1439

DB 487 CTGTGCTGAGCACTAATGCGACTATCACTCCAGGAGAGAAATTTAAATGGATG 546
QY 1440 TTGCGCGCAGCAGAGAGTATTTCTATGAATGTGGGCTTTTGGGCAAGGCAATAGCTCT 1499
DB 547 CTCAGAGCGGCTGAAGATATTTCTATGNACTTAGGTTTCTTGGCTAAAGGTAAACGCTTCT 606
QY 1500 AGCAAAAAACAATTTGTAGAACAAAGTAGAACGGGCGCGATTTGTTTTAAATTCATGAA 1559
DB 607 AACGACGCGAGCTTAGCCGATCAAAATTTGAAGCTGGTGGATTTGCTTTAAATCCACGAA 666
QY 1560 GACTGGGCGACAAACCAAGTCGATGCTGCTTGGAGCGTGGCAGATGAATACGAT 1619
DB 667 GACTGGGCGACCACTCTCTTCTGCAATCAATCATGCGTTAGATGTTGCGACAAATACGAT 726
QY 1620 GTGCAAGTTTGTATCCACACCGGATACAGTCAATGAGGCGAGTTATGTAGATGACACCCCTA 1679
DB 727 GTGCAAGTCTGCTATCCACACAGACATTTGAAAGCGGTTGCGTGGGAAGACACTATG 786
QY 1680 AATGCAATGAACGGGCGGCGCATTCATGCGCTACACATTTAGGGAGCGGGTGGAGGACAC 1739
DB 787 GCAGCTATTGCGCGAGCGCACTATGCACACTTTCACACTGAAGGTGCTGGCGGGGACAC 846
QY 1740 TCACCTGATGTTATCACTATGCGAGCGAGGTCAATTTCTACCTCTCTCACACACCCCT 1799
DB 847 GCTCTGATATTTAAAGTAGCTGGTGAACACAACTTTCTCCGCTTCCACTAACCCCT 906
QY 1800 ACTATTCCTTATACCATTAATACGTTTGCAAGACACTTAGACATGCTCATGACATGCCAC 1859
DB 907 ACTATTCCTTATGTAATACAGAACGAGACACATGGACATGCTTATGGTGTGCCAC 966
QY 1860 CACTAGACAAACGCACTCCGCGAGGATTTCAATTTTCTCAAAGCGGTATCCGCCCGGC 1919
DB 967 CACTTGGATAAAGCAATTAAGAAGATGTTTCAGTTCGCTGATTCGAAGGATCCGCCCTCAA 1026
QY 1920 TCTATCGCGCTGAAGATGCTCCATGATGATGGGTGATGCGGATGACAAAGCTCGGAT 1979
DB 1027 ACCATTTGCGGCTGAAGACACTTTGCTATGATGGGATTTTCTCAATCACCAGCTCTGAC 1086
QY 1980 TCCAGCAATGGGCGTGCAGGCGAAGTATCTCTCGAATCTGCGACACTGCGGATAG 2039
DB 1087 TCTCAAGCTATGGGTGCTGGGGTGAAGTTATCTAGAACTTGGCAACAGCTGACAAA 1146
QY 2040 AATAAAAAAGAAATTTGGTAAAGTTCCTGAAAGATGGCAAGATAACGATAATTTCCGCAAT 2099
DB 1147 AACAAAAAGAAATTTGGCGCTTGAAGAGAGAAAGCGGATACGACAACTTCAGGATC 1206
QY 2100 AAGCGCTACATCTCAAATAACATCATCAACCCCGCTTTGACCCAGCGGTGAGCGAGTAT 2159
DB 1207 AAAGCTTACTTGTCTAAATACACCATTAACCCGCGATCGCTCATGGGATTAGCGAGTAT 1266
QY 2160 ATCGGCTCTGTGGAAGGCGAAGATCGCGACTTGGTGGTGGGATCTGCTCTTTT 2219
DB 1267 GTAGGTTCTGTAGAGTGGGCAAGTGGCTGACTTGGTATTTGGAGTCCCGCATTTCTTT 1326
QY 2220 GGGTAAAAACCCAAATTCGTGATCAAAAGCGGTATGGTGGTCTCTCTGAAATGGCGCAT 2279
DB 1327 GGGTAAAAACCAACATGATCATCAAGGCGGGTTCATTTGGGTTGAGTCAATGGGTGAC 1386
QY 2280 TCTAACCGCTGTGTGCCCACTCCCAACCGGTTTATTTACCGCGAAATGTTTGGGATCAC 2339
DB 1387 GCGAACGCTTCTATCCCTACCCCAACCAAGTATTTATACAGAGAAATGTTGCTCATCAT 1446
QY 2340 GGCAGGCGAAATTTGACACACGATCATCTTTTGTTCGAAAGTGGCTATGAAAATGGC 2399
DB 1447 GGTAAAGCCAAATACGATGCAAAACATCACTTTTGTCTCAAGCGGCTTATGACAAAGC 1506
QY 2400 GTGAAGAAAGCTGGGCTTAGCGCGCAAGTCTTACCGGTCAAAAACCTGCGGTAACTC 2459
DB 1507 ATTAAGAAAGAAATTTAGGCGCTTGAAGAGCAAGTGTGGCGGTAAAAAATTTGCAAAACATC 1566
QY 2460 ACCAAGAAAGACTTCAAGTTCAACGACAAAAACGCAAAAAATCACCGTCGATCCGAAAAAC 2519
DB 1567 ACTAAAAAGACATGCAATTCACGACACTACCGCTCACAATGAAGTCAATCTGAACT 1626

QY 2520 TTCGAGGTCTTTGTAGATGCCAAACTCTGCACCTCTTAAACCCACCTCGCAAGTGCCTCTA 2579
Db 1627 TACCATGTGTTCGTGGATGCCAAAGAGTAACCTTCTAAACGAGCAATAAAGTGAGCTTG 1686
QY 2580 GCCAGGCTACACTTCTCTCTAG 2603
Db 1687 GCGCAACTCTTTAGCATTTTCTAG 1710

RESULT 12

US-10-500-447A-5
; Sequence S, Application US/10500447A
; Publication No. US20050150016A1
; GENERAL INFORMATION:
; APPLICANT: PARK, Hee-Sung
; TITLE OF INVENTION: Method for producing a recombinant protein using pollen
; FILE REFERENCE: YLOP040518US/PCT
; CURRENT APPLICATION NUMBER: US/10/500,447A
; CURRENT FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: KR 2001-71712
; PRIOR FILING DATE: 2001-11-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 1710
; TYPE: DNA
; ORGANISM: Helicobacter pylori
US-10-500-447A-5

Query Match 30.5%; Score 880; DB 9; Length 1710;
Best Local Similarity 69.8%; Pred. No. 1.5e-190;
Matches 1189; Conservative 0; Mismatches 515; Indels 0; Gaps 0;
QY 900 AAAATGAAAAAACAAGATATGTAATACTACGAGCCACCAAGGCGATAAAGTGGC 959
Db 7 AAGATTAGCAGAAAAAGAAATATTTCTATGTATGGCCCTACTACAGGCGATAAAGTGAGA 66

QY 960 TTAGGAGATACCGATCTTTGGGCGAGTAGACATGACTATACCACCTATGGCGAGAA 1019
Db 67 TTGGGCGATACAGACTTGATCGTGAAGTAGACATGACTACACCACTTTATGGCGAAGAG 126

QY 1020 CTTAAATTTGGCGGGTAAAACTATCCGTGAGGTATGGTCAGAGCAATAGCCCTGAT 1079
Db 127 CTTAAATTTGGCGGGTAAAACTTAAAGAGGCAATAGCCATTAACAACCTTAGC 186

QY 1080 GAAAAACCCCTAGATTAGTTCATCACTAACCGGATGATTATGCACTACACCGGATTTAC 1139
Db 187 AAAGAAGAACTGGATCTAATCATCACTAACCGCTTTAATCGTGGATTACACCGGTATTTAT 246

QY 1140 AAAGCCGACATTTGGGATTTAAACCGGCAAAATCCATGGCAATGGCAAGGAGGAACAAG 1199
Db 247 AAAGCGGATATTGGTATTAAAGATGGCAAAATCGCTGGCATTTGTTAAAGCGGTAAACAAA 306

QY 1200 GACATCGAGATGCGGTAAAGCCCTCATATGGTGGTGGTGGGCAAGAGCACTAGCA 1259
Db 307 GACATCGAAGATGCGGTAAAGCAATCTTAGCGTGGGTCTGCTACTGAAGCCTTAGCC 366

QY 1500 AGCAAAAAACAACCTTGTAGAAACAAGTAGAGCGGCGCGATTTGGTTTAAATTTGCATGAA 1559
Db 607 AACGATCAAGCTTTAGCCGATCAAATTTGAAGCCGCTGCGATTTGGCTTTAAAAATCCACGAA 666
QY 1560 GACTGGGGCAACAACCAAGTGGCATCGATCACTGCTTTGAGCGTGGCAGATGAATACGAT 1619
Db 667 GACTGGGGCAACCACTCTCTTCTGCAATCAATCATGCGTTAGATGTTGCGGACAAATACGAT 726
QY 1620 GTGCAAGTTTGTATCCACACCGATACAGTCAATGAGGCGAGTTATGTAGATGACACCCCTA 1679
Db 727 GTGCAAGTTCGCTATCCACACAGACACTTTTGAATGAAGCGGTTGTGTAGAGACACTATG 786
QY 1680 AATGCAATGAACGGCGCGCCATCCATGCCCTACCACTTTGAGGGAGCGGTTGGAGACAC 1739
Db 787 GCAGCCATTTGCGGAGCGCACTATGCACACTTTCCACACTGAAGGCGCTGGTGGCGACAC 846
QY 1740 TCACCTGATTTATCACCATGGCAGGCGAGCTCAATATTCTACCTCTCCACCAACCCCTC 1799
Db 847 GCTCCTGATATTATTAAGTAGCTGGTGAACACACACTTCTGCCGCTTCCACTAACCCCTC 906
QY 1800 ACTATTCCCTATACCATTAATACGTTGCAGAACACTTAGACATGCTCATGACATGCCAC 1859
Db 907 ACTATCCCTTTCACTGTGAATACAGACGACACATGGAATGCTTATGTGTGCCAC 966
QY 1860 CACCTAGACAAACGCACTCCGCGAGGATTTTCAATTTTCTCAAAGCGGTATCCGCCCCGCGC 1919
Db 967 CACTTGGATAAAGCATTTAAGAAGATGTTTCACTTCGCTGATTCAGGATCCGCCCTCAA 1026
QY 1920 TCTATCCGCGCTGAAGATGTGCTCATGATATGGTGTGATCGCATGACAAAGCTCGGAT 1979
Db 1027 ACTATTGCGGCTGAAGACACTTTGCATGACATGGGATTTTCTCAATCACCAGTTCTGAC 1086
QY 1980 TCGCAAGCAATGGGCGGTGACAGGCAAGTGAATCTCGAACTTGGCAGACTGCGGATAG 2039
Db 1087 TCTCAAGCTATGGTCTGTTGGTGAAGTATCACCAAGACTTGGCAACAGCTGACAAA 1146
QY 2040 AATAAAAAGAAATTTGGTAAGCTTCTCAAGATGGCAAGATAACGATAATTTCCGCAATT 2099
Db 1147 AACAAAAAGAAATTTGGCCGCTTGAAGAGAAAGAGCGATTAACGACAACTTCAGGATC 1206
QY 2100 AAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACCCAGCGGTGAGCGAGTAT 2159
Db 1207 AAACGCTACTTGTCTAAATACACCATTAACCCAGCGATCGCTCATGGGATTAGCGAGTAT 1266
QY 2160 ATCGGCTCTGTGGAAGGCGCAAGATCCCGACCTTGGTGGTGGGAAATCTCGCTCTTTT 2219
Db 1267 GTAGGTTCTGTAGAAGTGGGCAAGTGGCTGACTTGGTGGTGGTGGAGTCCCGCATTTCTTT 1326
QY 2220 GGCGTAAACCCCAAAATCGTGATCAAAAGCGGTATGGTGGTCTTCTCTGAAATGGCGAT 2279
Db 1327 GGCGTAAACCCCAAAATCGTGATCAAAAGCGGTATCATTAAGCGGATTCATTTAGTCAAAATGGGTAT 1386
QY 2280 TCTAAACCGCTGTGTGCCCACTCCCAACCGGTTTATTAACCGGAAATGTTTGGGCGATCAC 2339
Db 1387 GCGAAGCGCTTCTATCCCTACCCCAACCGGTTTATTTATAGAGAAATGTTGCTCATCAT 1446
QY 2340 GGCAAGCGGAAATTTGACACCAAGCATCACTTTTGTTCCAAAGTGCCTATGAAATGGC 2399
Db 1447 GGTAAAGCTAAATACGATGCAAAACATCACTTTTGTCTCAAGCGGCTTATGACAAAGC 1506
QY 2400 GTGAAAGAAAGCTGGGCTTAGAGCGCAAGTTCTACCGGTCAAAAACCTGCGGTAAACATC 2459
Db 1507 ATTAAGAAAGAAATTTAGGCGTTGAAGGCAAGTGTGGCGGTAAAAAATTTGCAAAACATC 1566
QY 2460 ACCAAGAAAGACTTTCAAGTTCAACGACAAAAACGCAAAAATCACCGTGCATCCGAAAAAC 2519
Db 1567 ACTAAAAAGACATGCAATCAACGACACTACCGTGCATTTGAAGTCAAATCTGAACT 1626
QY 2520 TTCGAGTCTTTGTAGATGGCAAACTCTGCACTCTTAAACCACTCTGCAAGTGCCTCTA 2579
Db 1627 TACCATGTGTTGTTGGATGGCAAGAGTAACCTTCTAAACCAAGCAATAAAGTGAGCTTG 1686

| | | | |
|---|------|--|------|
| Qy | 2580 | GCCAGCGCTACACTTTCCTCTAG | 2603 |
| | | | |
| Db | 1687 | GCACAACCTTTTAGCAATTTCTAG | 1710 |
| | | | |
| RESULT 13 | | | |
| US-10-476-313-11 | | | |
| ; Sequence 11, Application US/10476313 | | | |
| ; Publication No. US20040241175A1 | | | |
| ; GENERAL INFORMATION: | | | |
| ; APPLICANT: MCKENZIE, BRENT | | | |
| ; APPLICANT: BOYLE, JEFFREY | | | |
| ; APPLICANT: LEW, ANDREW | | | |
| ; TITLE OF INVENTION: Antigen Targeting | | | |
| ; FILE REFERENCE: BDWP-002 | | | |
| ; CURRENT APPLICATION NUMBER: US/10/476,313 | | | |
| ; CURRENT FILING DATE: 2003-10-24 | | | |
| ; PRIOR APPLICATION NUMBER: PRS241 | | | |
| ; PRIOR FILING DATE: 2001-05-25 | | | |
| ; NUMBER OF SEQ ID NOS: 12 | | | |
| ; SOFTWARE: Patentln version 3.1 | | | |
| ; SEQ ID NO 11 | | | |
| ; LENGTH: 1717 | | | |
| ; TYPE: DNA | | | |
| ; ORGANISM: Helicobacter pylori | | | |
| US-10-476-313-11 | | | |
| Query Match 30.4%; Score 876.8; DB 8; Length 1717; | | | |
| Best Local Similarity 69.7%; Pred. No. 8.1e-190; | | | |
| Matches 1187; Conservative 0; Mismatches 517; Indels 0; Gaps 0; | | | |
| Qy | 900 | AAAAATGAAAAAACAGAATAATGTAAATACCTACGGACCACCAAAGCGGATAAAGTGCGC | 959 |
| Db | 13 | AAGATTAGCAGAAAAGAATAATGTTCTTATGTATGCCCTACTACAGGTGATAAAGTGAGA | 72 |
| Qy | 960 | TTAGGAGATACCGATCTTTGGGAGAGATAGAACATGACTATACCACCTATGGCGAAGAA | 1019 |
| Db | 73 | TTGGCGGATACAGACTTGTATCGCTGGAAGTAGAACATGACTACACCATTTATGGCGAAGAG | 1332 |
| Qy | 1020 | CTTAATTTTGGCGGGTAAACATCTCCGTGAGGTTATGGGTACAGCAATAGCCCCTGAT | 1079 |
| Db | 133 | CTTAATTCGTGGCGGTAAACCCCTAAGAGAAGCATAGCCNACTTAACAACCCCTAGC | 192 |
| Qy | 1080 | GAATAACACCTAGATTTAGTTCATCACTAAACCGCATGATTATCGACTACACCGGGATTAC | 1139 |
| Db | 193 | AAAGAAGAAGCTGGATCTAAATCATCACTAAACGTTTTAATCGTGGATTACACCGGTATTTAT | 252 |
| Qy | 1140 | AAAGCCGACATTTGGATTTAAAAACGGGCAAAATCCATCGCATTTGGCAAGGCGAGAAACAAG | 1199 |
| Db | 253 | AAAGCGGATATTGGTATTAAAGATGGCAAAATCGCTGGCATTTGTTAAGGCGGTAAACAA | 312 |
| Qy | 1200 | GACATGCAAGATGGCGTAAGCCCTCATATGGTCGTGGGTGGGSCACAGAACACTAGCA | 1259 |
| Db | 313 | GACATGCAAGATGGCGTTAAAAACAATCTTAGCGTGGGTCTCTACTGTAAGCCCTTAGCC | 372 |
| Qy | 1260 | GGGGAAGGTATGATTATTACCGCTGGGGGAATCGATTACACACCCACTTCTCTTCTCCA | 1319 |
| Db | 373 | GGTGAAGGTTTGATCGTAACTGCTGGTGTATTGACACACACATCCACTTCATTTCAACC | 432 |
| Qy | 1320 | CACAAATTCCTTACCGCTCTAGCCAAATGGCGGTTAAACCATGTTTGGAGGCGGCACAGGT | 1379 |
| Db | 433 | CAACAAATTCCTTACAGCTTTTGGCAAGCGGTGTAAACAACCATGATTGGTGGCGGAATCTGGT | 492 |
| Qy | 1380 | CCTGTAGATGCGACGAATGGCACCTACTATCACTCCGGGCAAAATGGAACCTTGCACCGCATG | 1439 |
| Db | 493 | CCTGCTGATGGCACTTAAGCGCACTACTATCACTCCAGGTGAAGAAATTTAAATATGGATG | 552 |
| Qy | 1440 | TTGGCGCGCAGCAGAGATATTCTATGAATGTGGGCTTTTTTGGGCAAGGCAATAGCTCT | 1499 |
| Db | 553 | CTCAGAGCGGCTGAAGATATTCTATGAATTTAGGTTTCTTGGCTAAGGTAAAGCTTCT | 612 |
| Qy | 1500 | AGCAAAAAACAACCTTGTAGAAACAATAGAACGGGCGGATTTGGTTTAAATTCATGAA | 1559 |

RESULT 14
US-09-402-100-1
; Sequence 1, Application US/09402100
; Patent No. US20010019834A1
; GENERAL INFORMATION:
; APPLICANT: Daewoong Pharmaceutical Co., Ltd
; APPLICANT: Kim, Byung-O
; APPLICANT: Shin, Sung-Seup
; APPLICANT: Yu, Young-Hyo
; APPLICANT: Park, Myung-Hwan
; APPLICANT: Choi, Deok-Joon
; APPLICANT: Jung, Hyung-Jin
; TITLE OF INVENTION: Recombinant Microorganisms Expressing Antigenic Proteins of Helio
; FILE REFERENCE: 0136/0G140
; CURRENT APPLICATION NUMBER: US/09/402,100
; CURRENT FILING DATE: 1999-09-27
; EARLIER APPLICATION NUMBER: KR 97-11950
; EARLIER FILING DATE: 1997-03-31
; EARLIER APPLICATION NUMBER: KR 97-11951
; EARLIER FILING DATE: 1997-03-31
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2385
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1..1)
; OTHER INFORMATION: Recombinant DNA
US-09-402-100-1

Query Match 29.7%; Score 855.6; DB 3; Length 2385;
Best Local Similarity 69.8%; Pred. No. 6.9e-185;
Matches 1169; Conservative 0; Mismatches 504; Indels 1; Gaps 1;
QY 900 A A A A T C A A A A A C A A G A A T G T A A A T A C C T A C G G A C C C A C C A A A G G C G A T A A G T G C G C 959
DB 7 A G A T T A G C A G A A A A G A A T G C T T C T A T G T A T G G C C C T A C T A C A G C G C G A T A A A G T G A G 66
QY 960 T T A G G A G A T A C C G A T C T T T G G G C A G A A G T A G A A C A T G A C T A T A C C A C C T A T G G C G A A G A A 1019
DB 67 T T G G G C G A T A C A G A C T T G A T C G T G A A G T A G A C A T G A C T A C A C C A T T A T G T G T A A G A G 126
QY 1020 C T T A A A T T T G C G C G G T A A A A C T A T C C G T A G G G T A T G G G T C A G A G C A A T A G C C C T G A T 1079
DB 127 C T T A A A T T C G G C G G T A A A A C C C T A A G A G A A G C A T G A G C C A A T C T A A C A A C C C T A G C 186
QY 1080 G A A A C A C C T A G A T T T A G T C A T C A C T A A G C G A T G A T T A T C A C T A C A C C G G A T T T A C 1139
DB 187 A A A G A A G A A C T G G A T C T A A T C A T C A T A C G C T T T A A T C G T G A T T A C A C C G T A T T T A T 246
QY 1140 A A A G C G A C A T T G G G A T T A A A A C G C A A A T C C A T G G C A T T G C C A A G C A G G A A C A A G 1199
DB 247 A A A G C C G A T T T G T A T T A A G A T G G C A A A T C G T G G C A T T G T A A A G C C G T A C A A A 306
QY 1200 G A C A T G C A A G A T G C C T A A G C C C T A T A T G G T C G T G G G T G T G G C A C A G A A G C A T A G C A 1259
DB 307 G A C A C G A A G A T G C G T T A A A A C A A T C T A G C G T G G G T C C T A C T A G A A G C C T T A G C C 366
QY 1260 G G G A A G G T A T A T T A T T A C C G T G G G G A A T C G A T T C A C A C C A C C A C T T C C T T T C C A 1319
DB 367 G G T A A G G T T T G A T T G A A C T G C T G T G T A T T G A C A C A C A C A C A T C C A C T T C A T C T C C C C 426
QY 1320 C A C A A T T C C C T A C C C C T C T A G C C A T G G C G T T A C A C C A T G T T G A G C G C G A C A G G T 1379
DB 427 C A A C A A T C C C T A C A C G C T T T G A A G C G G T A A C A C C A T G A T T G G T G C G C G A A C T G G C 486
QY 1380 C C T G T A G A T G G C A A T G C A C T A C T A C C T C C G G G C A A A T G G A A C T T G C A C C G C A T G 1439
DB 487 C T G C T G A T G G C A C T A A C G C A C C A T A C T A C T C C A G G T A G A A A A T T T A A A A T T C A T G 546

QY 1440 T T G C G C G C A G C A G A A G A G A T T C T A T G A A T G T G G C T T T T T T G G C A A A G G C A A T A G C T C T 1499
DB 547 C T C A G A G C G C T G A A G A A T A T T C T A T G A A C T T T G G T T T C T T G G C T A A A A G G T A A C C T C T 606
QY 1500 A G C A A A A A C A A C T T G T A G A A C A A G T A G A A G C G C G A T T G G T T T T A A A A T T G C A T G A A 1559
DB 607 A A C G A T G C A A G C T T A G C C G A T C A A A T T G A A C T G T G C G A T T G C C T T A A A A T C C A C G A A 666
QY 1560 G A C T G G G C A C A A C A C C A A G T G C G A T C A C T C A C T T T G A G C G T G C G A G A T G A A T A C G A T 1619
DB 667 G A C T G G G C A C C A C T C C T T C T G C A A T C A A T C A T G C G T T A G A T T G T G C G A C A A A T A C G A T 726
QY 1620 G T G C A A G T T G T A T C C A C A C G A T A C A G T C A A T A G A G C A G G T A T G T A G A T G A C A C C T A 1679
DB 727 G T G C A A G T C G T A T C C A C A C A G A C A C T T T G A A T G A A G C C G T T G C G T G G A G A C A C A T A G 786
QY 1680 A A T G C A A T G A A C G G C G C G C C A T C C A T C C T A C C A C A T T G A G G A G C G G T G G A G A C A C 1739
DB 787 G C A G C T A T T G C C G A C G C A C T A T A G C A C A C T T A C C A C A C T G A A G G C G T G G G C G C G A C A C 846
QY 1740 T C A C C T G A T G T T A C A C C A T G C G A G C T C A A T A T T C T A C C C T C C T C C A C C A C C C C C 1799
DB 847 G C T C C T G A T A T T A A A G T G C C G T G A A C A C A C A C A T C C T A C C C G C T T C A C T A A C C C C 906
QY 1800 A C T A T T C C C T A T A C C A T T A A T A C G T T G C A G A A C A C T T A G A C A T G C T C A T G A C A T G C C A C 1859
DB 907 A C T A T C C C T T T C A C C G T G A A T A C A G A A G C C G A C A C A C A T G G A C A T G C T T A T G T G T G C C A C 966
QY 1860 C A C T A G A C A A A C G C A T C C G G A G G A T T T A C A A T T T C T C A A G C C G T A T C G C C C C G C 1919
DB 967 C A C T T G G A T A A A A G C A T T A A A G A A G A T G T C C A G T T C G C T G A T T C A A G A T T T G C C C T C A A 1026
QY 1920 T C T A T C G C G C T G A A G A T G T C C A T G A T A T G G T G T A T G C G A T G C A A G C T C G G A T 1979
DB 1027 A C C A T T C G C G C T G A A G A C A C T T T G C A T G A C A T G G G A T T T T C T A A T C A C T A G T T C T G A C 1086
QY 1980 T C G C A A C A A T G G G C G T G C A G C G A A G T A T T C T C G A A C T T G G C A G A C T G C G G A T A A G 2039
DB 1087 T C T C A G C A T G G C C G T G G G T G A A G T A T C A C T A G A A C T T G C A A C A C A G C T G A C A A A 1146
QY 2040 A A T A A A A A A A T T G T A A G C T T C T G A A G A T G G C A A A G A T A A C A T A A T T T C C G C A T T 2099
DB 1147 A A T A A A A A A A T T T G C C G C T T C A A A A A A A A A A A A G C G A T A A C A C A A C T T C A G A G A T C 1206
QY 2100 A A G G C T A C A T C T C C A A A T A C A C T A T C A A C C C C G T T G A C C C A G C G G T G A G C A G A T A T 2159
DB 1207 A A A C G C T A C T T G T C T A A A T A C A C C A T T A A C C C A G C G A T C G C T C A T G G G A T T A G C G A G A T 1266
QY 2160 A T C G G C T C T G T G A A G A G G C A A G A T C G C A C T T G G T G G T G T G G A A T C C T G C C T T T T T 2219
DB 1267 G T C G G T T C T G T A G A G T G G G C A A G T G C T G A C T T G T A T T G T G A G T C C C G C A T T C T T 1326
QY 2220 G G C G T A A A A C C C A A A T C G T G A T C A A A G C G G T A T G T G G T C T T C T C T G A A A T G G C G A T 2279
DB 1327 G G T G T G A A C C C A C A T G A T C A T C A A A G C G G T T C A T C G C A T T G A G T C A A A T G G T G A T 1386
QY 2280 T C T A A C G C G T G T G C C C A C T C C C C A C C G G T T T A T A C C G G A A A T G T T G G G A T C A C 2339
DB 1387 G C G A A C G C T T C T A T C C C C T A C C C C A C A A C C A G T T T A T A C A G A A A T G T T C G C T C A T C A T 1446
QY 2340 G G C A A G C G A A A T T T G A C A C C A G C A T C A C T T T T T C C A A A G T C C C C T A T G A A A T G C C 2399
DB 1447 G G T A A A G C T A A A T A C A G A T G C A A A C A T C A C T T T T G T G T C A A G C G G C T T A T G A C A A A G C 1506
QY 2400 G T C A A A A A A A G C T G G C T T A G A G C C C A A G T T C T A C C G G T C A A A A A C T G C C G T A A C A T C 2459
DB 1507 A T T A A A G A A A T T A G G G C T T G A A A G A C A A G T G T T G C G G T A A A A A A T T G C A A A A T A T C 1566
QY 2460 A C C A A A A A G A C T T C A A G T T C A A C G A C A A A A C G G C A A A A T C A C C G T C G A T C C G A A A C C 2519
DB 1567 A C T A A A A A A G A C A T G C A A T T C A A G C A C A C T A C C C G C T C A C A T T G A A G T C A A T T C T G A A A C T 1626

QY 2520 TTCGAGGTCTTTGTAGATGGCAAACTCTGCACCTCTTAAACCCACCTCGCAAGTG 2573
DB 1627 TACCATGTGTTCTGTGGATGGCAAA-GAAGTAAGTCTTAAACCGCAATATAAGTG 1679

RESULT 15

US-10-282-122A-32950
; Sequence 32950, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 32950
; LENGTH: 2341

; TYPE: DNA

; ORGANISM: Proteus mirabilis

US-10-282-122A-32950

Query Match 25.5%; Score 734.6; DB 7; Length 2341;
Best Local Similarity 58.8%; Pred. No. 3.4e-157;
Matches 1408; Conservative 0; Mismatches 929; Indels 66; Gaps 5;

QY 207 TGAAGCTCACACCAAGAGCAAGAAAAGTCTTGTATATATATGCGGGCGAAGTGGCTA 266
DB 2 TGGATTATACCAAGAGAAAAGATAATTAAGTCTTTTACTGAGGCTTGTTCAG 61
QY 267 GAAAGCGCAAGCAGAGGCTTAAAGCTCAACCAACCGAGCCATTGCTTACATTAGTG 326
DB 62 AAAGAGCTTTAGCTAAGGATTAAGCTTAATTAACCTGAAGCTGCGCTTGATTAGCT 121
QY 327 CCCATATTATGACGAAGCGCGCTGGAAAAAACCCTTGCCAGCTTATGAAGAGT 386
DB 122 GCGCCATTATGGAAGCGCACGAGAGG---GAAAAAGTGTGCTCAATTAATGAGTGAAG 178
QY 387 GCATGCATTTTGAAGAAAAGATGAAGTAATGCCGGGGTGGGTAATATATGTTCCCGATC 446

DB 179 GACGTACTGTTTAAACCGCAGAGCAAGTAATGGAAGGGGTGCCAGAGATGATAAAGATG 238
QY 447 TAGGTGTAGAAAGCCACCTTTTCCTGATGGTAACTTGTAACTGTGTAATTTGGCCCAATCG 506
DB 239 TTCAAGTAGAGTGCACTTTCCCGATGGCACCAAAATTTGGTTTCAATTCACCTACTTGG 298
QY 507 AACCATGAGCACTTCAAAGCGGGCGAAGTGAATAATTTGGTTGCGATAAAA---GACATCG 563
DB 299 TGT-AGGTAATAACATGATCCCGGTGAAATAGAGTTAATGCAGCATTAGGCGATATTG 357
QY 564 AGCTCAATGCAGCAAGAAAGTAACCGAATCTTGAGGTACTAATGAAGGGCTTAATCCT 623
DB 358 AACTGAATGCTGTCGCGAGACAAAAACCATACAGGTGGCTAATCATGGCGATAGACCTG 417
QY 624 TGCATGTGGGTAGCCATTTCCACTTTCTTGAAGCTAAACAGGCACTAAAAATTCGATCGTG 683
DB 418 TACAAGTCGGCTCTCATTAACCACTTTTGAAGTTAATGAGGCACTCAGGTTTGCACGAG 477
QY 684 AAAAAAGCTATGGCAACGCTAGATATTCCTCTGGCAACACGCTACGCAATTTGGGGCAG 743
DB 478 AAGAGACATTTAGGTTTTTCGTTTAAATATTCCTGCTGTATGGCTGTCGCTTCGAGCCG 537
QY 744 GACAAACCGCAAGTGCAGTTGATTCCTCTTGGTGGCAGTAAAAAGTGAATGGCATGA 803
DB 538 GTCAAGCGCGCACTGTTGAGTTAGTGGCTTTTGCAGGAAAAACGTGAAATTTATGG- 592
QY 804 ACGGGCTGTGTAATAACATCGCGATGAACGCCATAAACAATAAAGCGCTTCACAAAGCGA 863
DB 593 -----TTTTTCATGGCAA 604
QY 864 AATCTCACGGATTTATCAAGTAAGGAGACTCCCATGAAAAATGAAAAACAAGAAATATGTA 923
DB 605 AGTGATGGGTAATTCGAGAGTGAGAAAAAATAAGAACTATCTCACGCTCAAGCTTATGCG 664
QY 924 AATACCTACGACCCCAAGCGGATAAAGTGCGCTTAGGAGATACCGATCTTTGGGCA 983
DB 665 GATATGTTTGGCCCAACAACAGCGCATCGTTTGGATTTAGCAGATACCGAGCTGTTTCTT 724
QY 984 GAAGTAGAACATGACTATACCTATGCGCAAGAACTTAAATTTGGCGGGTAAAACT 1043
DB 725 GAAATTTGAAAAAGATTTTCAACCTTATGGCAAGAGGTCAAATTTGGTGGTGGTAAAGTT 784
QY 1044 ATCCGTGAGGGTATGGGTACAGCAATAGCCCTCATGAAAAACACCCCTAGATTTAGTCATC 1103
DB 785 ATTCGTGATGTTATGGGCAAGCCAAAGTTGTTAGTCTGAGTGTGTCGATGTTCTGATC 844
QY 1104 ACTAACGCGATGATTATCGACTACACCGGATTTACAAAGCGGACATTTGGGATTAAGAAC 1163
DB 845 ACCAATGCCATTTATTTAGATTTATTTGGGCAATTGTAAGAGCAGATATTGGCATTTAAGAT 904
QY 1164 GGCAAAATCCATGGCATTTGCAAGGCAAGGAACAAGGACATGCAAGATGGCGTAAAGCCT 1223
DB 905 GGCCGTATTGTCGGTATTGGCAAGGCGAGGTAATCCAGATGTTTCAGCCCAATGT-----G 958
QY 1224 CATATGTCGTGGGTGGGCACAGAAGCACTAGCAGGGGAAGGTATGATTATTACCGCT 1283
DB 959 GATATTGTCAATTGGCCCGGAAACAGAAAGTTGTCGGGGAGAGGTAAATATGATCACTGCT 1018
QY 1284 GGGGGAATCGATTACACACCCACTTCTTTCTCCACAACAATTTCCCTACCGCTTAGCC 1343
DB 1019 GGTGGTATTGATACCCATATCCACTTTTATTTGTCACAAAGCCCAAGAGGTCTCGTT 1078
QY 1344 AATGGCGTTACACCATGTTTGGAGGCGGACAGGTCCTGTAGTGGCACGAATCGGACT 1403
DB 1079 TCTGGCGTAACCACTTTATTTGGTGGAGGAACAGGCCCTGTGGCAGGTACTAATGCAACC 1138
QY 1404 ACTATCACTCGGGCAAAATGGAACCTGCACCGCATGTTGCGCGCAGCAGAGAGATTCT 1463
DB 1139 ACAGTTACCCCGGTATTTGGAATATGTACCGCATGTTAGAGCGGTGATGAATACCT 1198
QY 1464 ATGAATGTGGCTTTTTTGGGCAAGGCAATAGCTCTTAGCAAAAAACAATTTGTAGAACA 1523
DB 1199 ATTAAATGTGGGTTATTTGGTAAAGGTTGTGTGTCAGTCAGCCCGAAGCAATCCGCGACA 1258

| | | | |
|----|------|--|------|
| Qy | 1524 | GTAGAAAGCGCGCGAATTGGTTTTAAATTCATGAAGACTGGGGCAACAACCAAGTGGC | 1583 |
| Db | 1259 | ATAACAGCGGGTGTATAGTCTTAAAAATACATGAAGACTGGGGGCAACGCCAATGGCA | 1318 |
| Qy | 1584 | ATCGATCAGTCTGTAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACCGAT | 1643 |
| Db | 1319 | ATTCAAAATTGGCTTAATGTGCGCGATGAATGGATGTACAGGTGGCTATTCATCTGAC | 1378 |
| Qy | 1644 | ACAGTCAATGAGCGAGTTATGTAGATGACACCTTAAATGCAATGAACGGCGCGGCATC | 1703 |
| Db | 1379 | ACCTTAAATGAAGTGGTTTTTATGAAGACAGATGAAGCCATTTGCCGTCGAGTGATC | 1438 |
| Qy | 1704 | CATGCCCTACCAATTTGAGGAGCGGGTGGAGGACACTCACCTGATGTATCACCATGGCA | 1763 |
| Db | 1439 | CATGTATTCATACGACGAGCGAGTGGTGGTTCATGCCCTCATGTGATCAAGTCGGTA | 1498 |
| Qy | 1764 | GGGAGCTCAATATTTCTACCTCTCCACCAACCCCACTATTCCTCTATACCATTAATACG | 1823 |
| Db | 1499 | GGAGAGCCCAATATTTTACCTGCATCAACCAACCAACGATGCTTTATACCATTAATACC | 1558 |
| Qy | 1824 | GTTCAGACACCTTAGACATGCTCATGACATGCCACCACTAGACAAACGATCCGGAG | 1883 |
| Db | 1559 | GTGGACGAGCATCTTGATATGTGTGATGCTGTGCATCATCTCGATCCCTCATTTCTGAA | 1618 |
| Qy | 1884 | GATTTACAATTTTCTCAAAGCCGTATCCGCCCGGCTCTATCGCGCTGAAGATGTGCTC | 1943 |
| Db | 1619 | GATGTGGCAATTTGCTGAATCTGTATTCGTGCGGAACCACTTGTCTGCAGAGATATCTTA | 1678 |
| Qy | 1944 | CATGATATCGGTGTGATCGCGATGACAAAGCTCGGATTCGCAACAAATGGGGCGTCAGGC | 2003 |
| Db | 1679 | CATGATATGGGGCAATTTTCGGTGTGTGCTCAGACTCAAGCCATGGGACGAGTCGGA | 1738 |
| Qy | 2004 | GAAGTGATTCCTGGAATCTTGGCGAGACTCGGATAAGAAATAAAAGAAATTTGGTAAGCTT | 2063 |
| Db | 1739 | GAAGTATATCTTAGCGACTTTGGCAGTGTGCACATAAAATGAAATTTGCAACGAGGACATTA | 1798 |
| Qy | 2064 | CTGAGATGGCAAGACATCAACGATTAATTCGCGATTAAGCGCTACATCTCCAAATACACT | 2123 |
| Db | 1799 | GCGGTGATAGCGCAGATGAATGAATAATTCGTATTAACGTTATATCGGTAATAACAG | 1858 |
| Qy | 2124 | ATCAACCCCGCTTTTGACCCACGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCAAG | 2183 |
| Db | 1859 | ATTAAACCCGCACTGGCACATGGTATTTGCTCATACGGTGGATCAATAGAAAAGGTAAA | 1918 |
| Qy | 2184 | ATCGCCGACTTGGTGTGTGGAATCTCGCTTTTGGCGTAAACCCAAAATCGTGATC | 2243 |
| Db | 1919 | CTTGGGATATCGTGTATGGGATCCTGCTTTCTTTGGGCTCAACCGGCATTTATCATY | 1978 |
| Qy | 2244 | AAAGCGGTATGGTGTCTTCTGAAATGGGCGATTTCAACGGCTCTGTGCCACTCCC | 2303 |
| Db | 1979 | AAAGTGGTATGGTTCGCTTATGCGCAATGGGGGATATTAATCGCGCTATTTCCACACCG | 2038 |
| Qy | 2304 | CAACCGTTTATTACCGCAAAATGTTTGGGCATCAGCGAAGCGGAAAATTTGACCCAGC | 2363 |
| Db | 2039 | CAACCGTTTCAATATCGTCCAATGATGCTGTTTAGGAAAGCCAAATATCAACGTCG | 2098 |
| Qy | 2364 | ATCAGTTTTTTTCCAAAGTTCGCTATGAAAATGGCGTGAAGAAAAGCTGGGCTTAGAG | 2423 |
| Db | 2099 | ATGATCTTTATGTCAAAAGCGGTAATTTAGCGGGAGTGCAGAAAAATTTAGGCTTAAAA | 2158 |
| Qy | 2424 | CGCAAGTTCTACCGGTCAAAAACCTCGCGTAAACATCAACCAAGAAAGACTTCAAGTTCAAC | 2483 |
| Db | 2159 | AGCTTAATTTGGTGTGGAGGCTGTGCTCATATCACAAAAGCTTTCGATGATCCCAAT | 2218 |
| Qy | 2484 | GACAAAACGGCAAAATACCGTTCGATCCGAAACCTTCGAGGTCTTTGTAGATGGCAA | 2543 |
| Db | 2219 | AACATATGTTCTCATATCGAATTAGATCCCAAACTTACATTTTAAACGGGATGGTGA | 2278 |
| Qy | 2544 | CTCTGCACCTCTAAACCCCACTCGCAAGTGCCTCTAGCCCGAGCGCTACACTTTCTCTAG | 2603 |
| Db | 2279 | CCACTGGTTTGTGAGCCAGGACTGAATTAACCGATGGCTCAACGCTATTTCTTATTTAA | 2338 |

| | | | |
|----|------|-----|------|
| Qy | 2604 | GCA | 2606 |
| Db | 2339 | CCA | 2341 |

Search completed: November 29, 2005, 08:04:19
Job time : 2060 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2005, 00:34:45 ; Search time 563 Seconds
(without alignments)
760.989 Million cell updates/sec

Title: US-09-904-994B-1
Perfect score: 2883
Sequence: 1 rgragattttccarctt.....aaaaagtagagccacag 2883

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3205263 seqs, 74304013 residues

Total number of hits satisfying chosen parameters: 6410526

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.New.*
1: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2.*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3.*
10: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 612 | 21.2 | 3164 | 1 | US-10-793-626-3828 |
| 2 | 612 | 21.2 | 3234 | 1 | US-10-793-626-3853 |
| 3 | 578.6 | 20.1 | 1716 | 1 | US-10-793-626-117 |
| 4 | 455 | 15.8 | 1612 | 1 | US-10-485-517-26 |
| 5 | 424.4 | 14.7 | 3475 | 1 | US-10-793-626-4032 |
| C 6 | 216.2 | 7.5 | 526 | 1 | US-10-834-151-2 |
| C 7 | 97.4 | 3.4 | 309 | 1 | US-10-793-626-121 |
| 8 | 74.2 | 2.6 | 336 | 1 | US-10-793-626-115 |
| 9 | 46 | 1.6 | 2259 | 1 | US-10-793-626-4078 |
| C 10 | 40 | 1.4 | 3373 | 1 | US-10-793-626-4389 |
| C 11 | 40 | 1.4 | 3927 | 1 | US-10-793-626-3926 |
| C 12 | 39.8 | 1.4 | 3923 | 1 | US-10-793-626-4109 |
| C 13 | 37.8 | 1.3 | 1809 | 1 | US-10-793-626-2361 |
| C 14 | 37.8 | 1.3 | 3591 | 1 | US-10-793-626-4033 |
| C 15 | 37.2 | 1.3 | 2988 | 1 | US-10-793-626-4132 |
| C 16 | 37.2 | 1.3 | 4210 | 1 | US-10-793-626-3761 |
| C 17 | 37 | 1.3 | 3362 | 1 | US-10-793-626-3731 |
| C 18 | 37 | 1.3 | 4041 | 1 | US-10-793-626-3689 |
| C 19 | 36.2 | 1.3 | 738 | 1 | US-10-793-626-2909 |
| C 20 | 36.2 | 1.3 | 3218 | 1 | US-10-793-626-4229 |
| C 21 | 36.2 | 1.3 | 3997 | 1 | US-10-793-626-4220 |
| C 22 | 35.8 | 1.2 | 14941 | 1 | US-10-821-234-771 |
| C 23 | 35.6 | 1.2 | 1017 | 1 | US-10-793-626-1003 |

ALIGNMENTS

RESULT 1

US-10-793-626-3828
; Sequence 3828, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P0348005
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3828
; LENGTH: 3164
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3828

| | | | | |
|-----------------------|-----------------|--|------------|-------------------|
| Query Match | 21.2% | Score 612; | DB 1; | Length 3164; |
| Best Local Similarity | 57.4%; | Pred. No. 5.2e-139; | | |
| Matches 1211; | Conservative 0; | Mismatches 845; | Indels 52; | Gaps 4; |
| Qy | 543 | TTGGTTGCGATAAGACATCGAGCTCAATGTCAGGCAAGAGTAACCGAACTTTGAGTTA | 602 | Sequence 41, Appl |
| Db | 468 | TTGTTAAATAATCTGAATAAGTCAATAACATCATCCCGAAACGGTGTGAAGTAA | 527 | Sequence 56, Appl |
| Qy | 603 | CTAATGAAGGCGCTAAATTCCTTGCATGTGGGTAGCCATTTCCATCTTTTGAAGTAAACA | 662 | Sequence 3694, Ap |
| Db | 528 | AAATACAGGCGATAGACCTATACAAGTAGGTTCACATTTCCACTTTTCGAAGCAATA | 587 | Sequence 4222, Ap |
| Qy | 663 | AGGCACATAAATTCGATCGTGAAGAAAGCCTATGGCAACGCCCTAGATATTCCTCTGGCA | 722 | Sequence 4141, Ap |
| Db | 588 | AAGCATTAGATTTCGATCGTGAGAAAGCATATGGTAAACATTTGGATATTCCTGCGAGAG | 647 | Sequence 3, Appli |
| Qy | 723 | ACACGCTACGATTTGGGCGAGCAAAACCCGAGTCAGTTGATTCCTCTTGGTGGCA | 782 | Sequence 3919, Ap |
| Db | 648 | CTGACGATGAGATTGGAACCTGGAGATGAAAAAAGTAACAACCTGTCGATATTCCTGGAC | 707 | Sequence 3885, Ap |
| Qy | 783 | GTAATAAAGTGTATGGCATGAACGGGCTGTGTAATAAC- - -ATCGCGGATGAACGCCATA | 839 | Sequence 61, Appl |
| Db | 708 | GACGTAAATTTATGGATTCCGTGGTTAGTCGATGGCATATTGACGAAGAACGCGTAT | 767 | Sequence 3343, Ap |

| | | | | | | | | | | | | | | | | | | | | | | |
|----|------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|------|------|
| Qy | 840 | AA | CAT | AAG | CGC | TTG | ACA | AGCG | GAAT | CT | CA | CGG | ATTT | AT | CA | AGT | AAG | GAG | ACT | CCCA | TG | 899 |
| Db | 768 | TC | CGT | CAAA | TGAT | TC | AAA | T | CAAAA | CGCC | CGT | TT | AAAA | ACG | AT | GC | AGG | CGA | AGA | CAAT | G | 827 |
| Qy | 900 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 925 | |
| Db | 828 | CG | AA | T | AAAA | AAAG | TG | TT | AA | T | AG | CA | TG | AG | TT | T | AAAA | T | GAC | AA | CT | 887 |
| Qy | 926 | T | A | C | T | A | C | G | G | A | C | C | A | A | A | G | C | G | A | T | A | 985 |
| Db | 888 | T | C | T | T | A | T | G | G | A | C | A | A | C | T | G | T | A | G | A | T | 947 |
| Qy | 986 | A | G | T | A | G | A | A | C | A | T | A | T | A | C | C | T | A | G | C | G | 1045 |
| Db | 948 | A | G | T | T | A | A | A | A | A | G | A | C | T | A | T | A | G | A | T | A | 1007 |
| Qy | 1046 | C | G | T | A | G | G | T | A | T | C | G | G | T | C | A | G | A | T | A | T | 1096 |
| Db | 1008 | T | C | G | T | A | T | A | T | A | T | A | T | A | T | A | T | A | T | A | T | 1067 |
| Qy | 1097 | A | G | T | A | C | A | T | A | A | C | G | C | A | T | A | T | A | C | A | T | 1156 |
| Db | 1068 | A | G | T | T | A | A | C | T | A | A | T | A | T | A | T | A | T | A | T | A | 1127 |
| Qy | 1157 | T | A | A | A | A | C | G | G | A | A | A | T | C | C | A | T | G | C | A | T | 1216 |
| Db | 1128 | T | A | A | A | A | T | G | G | T | A | T | A | T | A | T | A | T | A | T | A | 1187 |
| Qy | 1217 | A | A | G | C | C | T | C | A | T | A | T | C | G | T | G | G | T | G | G | C | 1276 |
| Db | 1188 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 1241 | |
| Qy | 1277 | T | A | C | G | T | G | G | G | A | T | C | A | C | A | C | C | A | T | T | C | 1336 |
| Db | 1242 | T | A | C | T | G | C | G | C | G | T | A | T | C | G | A | T | C | A | T | T | 1301 |
| Qy | 1337 | T | C | T | A | G | C | A | A | T | T | C | A | A | C | C | A | T | G | T | T | 1396 |
| Db | 1302 | A | C | T | T | A | G | A | G | T | G | T | A | T | A | C | A | C | G | C | A | 1361 |
| Qy | 1397 | T | G | C | A | C | T | A | C | T | C | G | G | C | A | A | T | G | C | A | T | 1456 |
| Db | 1362 | A | G | C | A | C | T | T | G | A | C | A | C | A | C | A | T | T | C | A | T | 1421 |
| Qy | 1457 | G | T | A | T | T | C | T | A | G | A | T | G | G | C | A | A | T | A | G | C | 1516 |
| Db | 1422 | G | A | T | C | C | T | A | T | A | A | T | T | A | C | T | G | T | A | A | A | 1481 |
| Qy | 1517 | A | G | A | C | A | G | T | A | G | C | G | C | G | A | T | T | G | T | T | T | 1576 |
| Db | 1482 | T | G | A | A | C | A | A | T | T | C | A | T | G | C | T | A | A | G | A | T | 1541 |
| Qy | 1577 | A | A | G | T | G | C | A | T | C | A | T | T | G | A | C | G | T | G | A | T | 1636 |
| Db | 1542 | T | T | C | A | G | C | A | T | A | A | T | T | A | C | A | T | T | A | C | A | 1601 |
| Qy | 1637 | C | A | C | G | A | T | A | G | T | A | T | A | G | A | T | A | G | A | T | A | 1696 |
| Db | 1602 | T | G | C | A | G | A | T | A | A | T | A | A | T | A | A | T | A | A | T | A | 1661 |
| Qy | 1697 | C | G | C | A | T | C | C | A | T | T | A | G | G | A | G | C | G | G | T | A | 1756 |
| Db | 1662 | T | | | | | | | | | | | | | | | | | | | | |

RESULT 2

```

US-10-793-626-3853
; Sequence 3853, Application US/10793626
; Publication No. US20050255478A1
;
; GENERAL INFORMATION:
;
; APPLICANT: KIMMERLY, WILLIAM JOHN
;
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
;
; FILE REFERENCE: P03480US
;
; CURRENT APPLICATION NUMBER: US/10/793,626
;
; CURRENT FILING DATE: 2004-03-04
;
; PRIOR APPLICATION NUMBER: 60/164,258
;
; PRIOR FILING DATE: 1999-11-09
;
; NUMBER OF SEQ ID NOS: 4472
;
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 3853
;
; LENGTH: 3234
;
; TYPE: DNA
;
; ORGANISM: Artificial Sequence
;
; FEATURE:
;
; OTHER INFORMATION: Description of Artificial Sequence: synthetic

```

; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3853

Query Match 21.2%; Score 612; DB 1; Length 3234;
Best Local Similarity 57.4%; Pred. No. 5.3e-139;
Matches 1211; Conservative 0; Mismatches 845; Indels 5

| Query Match | 21.2% | Score 612 | DB 1 | Length 3234 |
|-----------------------|----------------|---|-----------|-------------|
| Best Local Similarity | 57.4% | Pred. No.5.3e-139 | | |
| Matches 1211 | Conservative 0 | Mismatches 185 | Indels 52 | Gaps 4 |
| Qy | 543 | TTGGTTGCGATAAAGACATCGAGCTCAATGCGAGCAAGAAAGTAACCGAACCTTCGAGGTAA | 602 | |
| Db | 537 | TTGTTAAATACTGAAATAGAGTCAATAAACAATCATCCGAAACGGTGATTGAAGTAA | 596 | |
| Qy | 603 | CTAATGAAGGGCCCTAAATCCCTTGTGATGTGGGTAGCCATTTTCCACTCTTTTGAAGCTAACA | 662 | |
| Db | 597 | AAAATAACAGGCGATAGACCTATACAAGTAGGTTCACATTTCCACTTTTTCGAAGCAATA | 656 | |
| Qy | 663 | AGGCACCTAAATTTGATCGTGTGAAAGAGCCCTATGCGAAACGCTAGATATATCCCTCTCGCA | 722 | |
| Db | 657 | AAGCATTTAGAAATTTGATCGTGAGAAAGCATATGTTAAACATTTTGGATATTTCTTCGAGGAG | 716 | |
| Qy | 723 | ACACGCTACGCATTTGGGCGAGGACAAACCCGCAAGTGCAGTTGATTTCCCTCTTGGTGGCA | 782 | |
| Db | 717 | CTGCAGTGAATTTGAACCTGGAGATGAAATAAAGTACAACCTTGTCTGAATATTTCTGGAC | 776 | |
| Qy | 783 | GTAATAAAGTGATTTGGCATGAAACGGGCTTTGTGAATAAC---ATCGCGGATGAACGCCATA | 839 | |
| Db | 777 | GACGTAAATTTTATGGATTTCGTTAGTCGATGGCGATATTTGACGAAGAACCGGTAT | 836 | |
| Qy | 840 | AACATAAAGCGCTTGACAGCGGAAATCTCACGGATTTTATCAAGTAGAGGAGACTCCCATG | 899 | |
| Db | 837 | TCCGTCCAAAATGATTTCAAAATCAAAACCCGCGCTTAAAAAACGATGTCAGGCGGAAGACAATG | 896 | |
| Qy | 900 | -----AAAATGAAAAAACAAGATATGTAAA | 925 | |
| Db | 897 | CGAATAAATAAGGTGGTAAATAAGCATGAGTTTTTAAATGACAAATCTCAATACACAAG | 956 | |
| Qy | 926 | TACCTACGGACCCACAAAGGCGATAAAGTCGCTTAGGAGATACCGATCTTTGGGCGAGA | 985 | |
| Db | 957 | TCCTTATGGACCAACTGTAGGAGACTCTGTGAGATTAGGAGATACGAACTTGTTTGCACA | 1016 | |
| Qy | 986 | AGTAGACATGACTATACCACTTATGCGAAGAACTTAAATTTGGCGGGGTAAACATAT | 1045 | |
| Db | 1017 | AGTTGAAAAAGACTATGCAAAATTTATGAGATGAAGCTACTTTCCGTCGCGGAAAAATCAAT | 1076 | |
| Qy | 1046 | CCGTGAGGGTATGGTGCAGA-----GCAATAGCCCTGATCAAAACACCCCTAGATTT | 1096 | |
| Db | 1077 | TCGTGATGGTATGGCTCAAATCTTAATGTGCAAGAGATGATAAAAATGTAGCCGATTT | 1136 | |
| Qy | 1097 | AGTCATCACTAACCGCATGATTCGACTACACCGGATTTACAAAGCCCGACATTTGGGAT | 1156 | |
| Db | 1137 | AGTTTTTAACTAACGCATTAATTAATGATTATGACAAGATTTTAAAGCAGATATCCGAAT | 1196 | |
| Qy | 1157 | TAAAAACGGCAAAATCCATGCGATTTGCGAAGGCGAGGAAACAAGACATGCAAGATGCGGT | 1216 | |
| Db | 1197 | TAAAAATGGTTATATTTTAAAGTCGGTAAAGCTGGAAACCCAGATATAATGGATAACGT | 1256 | |
| Qy | 1217 | AAGCCCTCATATGTCGTGGGTGGGCACAGAAGCACTAGCAGGGGAAGGTATGATTAT | 1276 | |
| Db | 1257 | -----TGACATCATCATTTGGTGCAACAACTGATATTTATTTGCTGCTGAAAGTAAATTTGT | 1310 | |
| Qy | 1277 | TACCGCTGGGGGAATTCGATTCACACACCCACTTCTCTTCTCCACAACAATTTCCCTACCGC | 1336 | |
| Db | 1311 | TACTGCCGGGGTATCGATACACACGTCGACTTTCATCAATCTCTGAAACAGCTGAAGTTGC | 1370 | |
| Qy | 1337 | TCTAGCCAAATGGCGTTACAACCATATGTTGAGGGCGGCACAGTCTCTGTAGATGCGACGAA | 1396 | |
| Db | 1371 | ACTTGAGATGGTATTAACAACGCATATCGGTGGAGGAACTGGTGCTTCTTGAAGGTGCTAA | 1430 | |
| Qy | 1397 | TGGCACTACTATCACTCCGGCAAAATGGAATCTCCACCGCATGTTGCGCGCAGCAGAAGA | 1456 | |
| Db | 1431 | AGCGACTACTGTAAACACAGAACCTTGGCATATTTTCATTCGCATGTTAGAACGACGAGAAGA | 1490 | |
| Qy | 1457 | GTATTTCTATGAATGTGGGCTTTTTTTGGGCAAAAGGCAATAGTCTTAGCAAAAAACAACCTTGT | 1516 | |

| | | | | |
|----|------|--|-------------------|------|
| Db | 1491 | GATGCTTATTAACTGTAGGATTTACTGGTAAAGGTCAAGCTGTCAATCATTA | CTGCACTTAT | 155 |
| Qy | 1517 | AGAAACAAGTAGAAGCGGCGCGATTTGGTTTTAAATTGCATGAAGACTGGGGCA | CAACACCC | 1576 |
| Db | 1551 | TGAACAATAATCATGAGGCGCTATAGTCTTAAAGTACATGAAGATTTGGGAGCT | ACACCC | 1610 |
| Qy | 1577 | AAGTGGCATCGATCACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTT | GTGTATCCA | 1636 |
| Db | 1611 | TTGAGCATTTAAGTCATGCAATTTAGACGTTTGAGATGAGTGTGATGTTC | CAAGTCGCTTTACA | 1670 |
| Qy | 1637 | CACCGATACAGTCAATGAGGCGAGGTTATGTAGTACACCTAAATGCAATGA | ACGCGCG | 1696 |
| Db | 1671 | TGCGAGATACATTAATGAAGCTGGATTTATGGAAGATACAATGGCTCTGT | GAAAGATCG | 1730 |
| Qy | 1697 | CGCCATCCATGCCCTACCACATTTAGGAGGCGGGTGGAGACACTCACCTG | ATGTTATCAC | 1756 |
| Db | 1731 | TGTATTGCATATGTATCATACTGAAGGCGTGGTGGTTCATGCACCTGACTT | AAATCAA | 1790 |
| Qy | 1757 | CATGGCAGGCGAGTCAATATTTACCCCTCTCCACCACCCCACTATTTCCCT | TATACCAT | 1816 |
| Db | 1791 | ATCAGCTGCATATTTCAAAACATCTTTACCTTCTTACAAACCCCAACTTA | CCCTTACACTCA | 1850 |
| Qy | 1817 | TAATACGGTTCAGAAACACTTTAGACATGCTCATGACATGCCACCCTAG | ACAAACGCAT | 1876 |
| Db | 1851 | CAACACTGTAGATGAACATTTAGACATGGTATTATTAATCTACCATCTT | TAATGCTTCAAT | 1910 |
| Qy | 1877 | CGCGAGGATTTACAAATTTTCTCAAAAGCGGTATCCGCGCGGCTCTATCG | CGGCTGAAGA | 1936 |
| Db | 1911 | ACGAGAAGACATTCGATTTGCGAGTTCTCGTATACGTAAAGAACTATAG | CAGCAGAGA | 1970 |
| Qy | 1937 | TGTGCTCCATGATATGGGTGTGATTCGCGATGACAAAGCTCGGATTCG | CAAGCAATGGGGCG | 1986 |
| Db | 1971 | CGTATTACAAGATATGGGCGTATTTAGTATGTTAAAGTTTCAGATTTCA | CAAGCAATGGGACG | 2030 |
| Qy | 1997 | TGCGAGGCGAAGTGATTCCTCGAACTTGGCAGACTCGCGATAAGATAA | AAAGAAATTTGG | 2056 |
| Db | 2031 | TGTCGGTGAAGTTGTAAACAGTACTTGGCAAGTTTGGCACACGATGA | AAAGAACACCGCG | 2090 |
| Qy | 2057 | TAAGCTCTTCAAGATGCGAAGATAACGATAATTTCCGCAATTAAGCGCT | TACATCTCAA | 2116 |
| Db | 2091 | ACCATTAGATGGTGACTTTGAAATATCAGATATAATCGTATTAAAGTT | ACATTTGCANA | 2150 |
| Qy | 2117 | ATACACTATCAACCCCGCTTTGACCCACGCGTGAAGAGTATATCGGCTCT | GTGGAAGA | 2176 |
| Db | 2151 | ATATACAATCAATCCTGCCATTACATGTTATTTCTGACTATGTTGGATCT | GTGTAGAAGC | 2210 |
| Qy | 2177 | GGGCAAGATCCCGACTTTGGTGTGTGAATCTCGCCTTTTGGCGTAA | AAACCCCAAT | 2236 |
| Db | 2211 | GGGTAAACTTGGCATTTTAGTAAATGTGGGAAACCAAGAAATTCCTCG | GTCACAAACCCGATCT | 2270 |
| Qy | 2237 | CGTGATCAAAAGCGGTATGGTGGTCTTCTGTAATGGCGATTTCTAAC | CGGCTCTGTGCC | 2296 |
| Db | 2271 | TGTTGTTAAAGGTGGCATGATTAATCAGCAGTAAATGGTGAATGCTAAT | TGGCTCCATACC | 2330 |
| Qy | 2297 | CACCTCCCAACCGGTTTATTAACCGCAAAATGTTTTGGCGATCACGG | CAAGCGCAATTTGA | 2356 |
| Db | 2331 | AACATCAGAGCTTTGAAATATCGCAAAATGTATGGTCAATTTTGGTGT | AAACATTTACACA | 2390 |
| Qy | 2357 | CACGAGCATCATTTTGTTTTCCAAAGTCGCCTATGAATAATGGCGTGA | AAAGAAAGCTGGG | 2416 |
| Db | 2391 | TACTGCTATGACTTTTGTCTTAACACTGCAATATGAATAATTTATTTAT | CGTCAACTCAA | 2450 |
| Qy | 2417 | CTTAGAGGCCAAGTCTACCGGTCAAAACCTGCGGTAAACATCACCA | AGAAAGACTTCAA | 2476 |
| Db | 2451 | TCTAAACGAATGGTTGCACAGTTAGAAATATTAGAAATTTAACTA | AGCGAGTATGAA | 2510 |
| Qy | 2477 | GTTCAACGACAAAAACGGCAAAATATCCCGTGCATCCGAAACCTT | TCGAGGCTTTGTAGA | 2536 |
| Db | 2511 | AAATAAATGCTACACCTAAATATAGTATAGATCCAAACATATGAGG | TATTCGTTGA | 2570 |
| Qy | 2537 | TGGCAAACTCTGCACCTCTAAACCACTCGCAAGTGCCTCTAGCC | ACGCGCTACATTT | 2586 |
| Db | 2571 | TGTTAATTAATAATCAAAAGTGAAGCAGCAACAGAAATTAACCATTA | TAAACAAGAGATCTTCT | 2630 |

| | | | | | | | |
|--|------|--|------|----|------|--|------|
| QY | 2597 | CTTCTAGG 2604 | | QY | 1489 | GCAATAGCTCTAGCAAAAAACAATTGTAGAACAGTAGAAGCGGCGCGATTTGGTTTTA | 1548 |
| Db | 2631 | ATTCTAGG 2638 | | Db | 602 | GTCAAGCTGTCAATCATCTATGTCACCTTATTGAACAAATTCATGACGGCGCTATAGTCTTTA | 661 |
| RESULT 3 | | | | QY | 1549 | AATTGCGATGAAGACTGGGCGACCAACCAAGTCGATCGATCACTGCTTGAGCGTGGCAG | 1608 |
| US-10-793-626-117 | | | | Db | 662 | AAGTACATGAAGATTGGGGAGCTACACCTTCAGCATTTAAGTCATGCAATTGACGTTGCGAG | 721 |
| ; Sequence 117, Application US/10793626 | | | | QY | 1609 | ATGAATACGATGTGCAAGTTTGTATCCACACCGATCAAGTCAATGAGGACGATTATGTAG | 1668 |
| ; Publication No. US20050255478A1 | | | | Db | 722 | ATGAGTTTGATGTTCAAGTCGCTTACATGACAGATACATTAATGAAGCTGGATTTATGG | 781 |
| ; GENERAL INFORMATION: | | | | QY | 1669 | ATGACACCCCTAAATGCAATGAACGGCGCGCATTCATGCCCTACACATGAGGAGCGG | 1728 |
| ; APPLICANT: KIMMERLY, WILLIAM JOHN | | | | Db | 782 | AAGATACAATGCGCTGCTGTAAGATCGTGTATTGTCATATGTATCATACTGAAAGGAGCTG | 841 |
| ; FILE REFERENCE: PU3480US | | | | QY | 1729 | GTGAGGACACTCACCTGATGTTATCACCATGGCAGCGAGCTCAATATTCTACCCCTCT | 1788 |
| ; CURRENT APPLICATION NUMBER: US/10/793,626 | | | | Db | 842 | GTGGTGTGTCATGCACCTGACTTAATCAAAATCAGCTGCAATATTCAAAACATCTTACCTTCT | 901 |
| ; CURRENT FILING DATE: 2004-03-04 | | | | QY | 1789 | CCACCAACCCCACTATTTCCTTATACCATTAATACGGTTGCAGAACACCTTTAGACATGCTCA | 1848 |
| ; PRIOR APPLICATION NUMBER: 60/164,258 | | | | Db | 902 | CTACAAACCCCACTTACCTTACACTCACACACTGTAGATGAACATTTAGACATGGTTA | 961 |
| ; PRIOR FILING DATE: 1999-11-09 | | | | QY | 1849 | TGACATGCCACCACTAGACAAACGCAATCCGCGAGGATTTTCAATTTTCTCAAAAGCCGTA | 1908 |
| ; NUMBER OF SEQ ID NOS: 4472 | | | | Db | 962 | TGATTACTCACCACTCTTAATGCTTCAATACAGAAAGACATTTGCATTTGACAGATTCGCTA | 1021 |
| ; SOFTWARE: PatentIn Ver. 2.1 | | | | QY | 1909 | TCGCCCCGGCTCTATCGCGCTGAAGATGTGCTCCATGATATGGTGTGATCGCGATGA | 1968 |
| ; SEQ ID NO 117 | | | | Db | 1022 | TACGTAAAGGAACTTATAGCAGCAGAAACGCTATTACAAGATATATGGCGCTATTAGTATGG | 1081 |
| ; LENGTH: 1716 | | | | QY | 1969 | CAAGCTCGGATTCGCAAGCAATGGGCGTGCGAGCGGAAGTATTCCTCGAACTTCGCGAGA | 2028 |
| ; TYPE: DNA | | | | Db | 1082 | TAAGTTTCAAGATTCACAAGCAATGGGACGTGTCGCTGAAGTTGTACACGCTACTTGGCAAG | 1141 |
| ; ORGANISM: Artificial Sequence | | | | QY | 2029 | CTCGGTAAGAATAAAAAAGAAATTTGGTAAGCTTCTCGAAGATTCCTGGAAGATGAACGATA | 2088 |
| ; FEATURE: | | | | Db | 1142 | TTGCACACCGGTATGAAAGAACACGCGGACCATTAGATGGTGACTTTTGAATATACGATA | 1201 |
| ; OTHER INFORMATION: Description of Artificial Sequence: synthetic | | | | QY | 2089 | ATTTCCGCATTAAGCGGTACATCTCCAAATACACTATCAACCCCGCTTTGACCCAGCGG | 2148 |
| ; OTHER INFORMATION: nucleic acid sequence | | | | Db | 1202 | ATAATCGTATTAAACGTTTACATTGCAAAATATACAATCAATCCTGCCATTACACATGATA | 1261 |
| US-10-793-626-117 | | | | QY | 2149 | TGAGCGAGTATATCGGCTCTGTGGAAGAGGCAAGATCGCCGACTTGTGTTGTGCGAATC | 2208 |
| Query Match 20.1%; Score 578.6; DB 1; Length 1716; | | | | Db | 1262 | TTTCTGACTATGTTGGATCTGTAGAGCGGTAACCTTGCCGATTTAGTAATGTGGGAAC | 1321 |
| Best Local Similarity 59.5%; Pred. No. 4.8e-131; | | | | QY | 2209 | CTGCTTTTGGCGTAAAAACCCAAAATCGTGTATCAAGGCGGTATGTTGTTCTTCTCTG | 2268 |
| Matches 1021; Conservative 0; Mismatches 679; Indels 15; Gaps 2; | | | | Db | 1322 | CAGAAATTTCTCGGTGCCAACCCCGATCTTGTGTTTAAAGTTGGCATGATTAACCTCAGCAG | 1381 |
| QY | 898 | TGAAATGAAACAAAGATATGTAATACCTACCGACCCCAACCAAGCGGATTAAGTGC | 957 | QY | 2269 | AAATGGCGGATTTCAACGCGTCTGTGCCCATCTCCCAACCGGTTTATTACCGCGAAATGT | 2328 |
| Db | 8 | TTAAATGACAAATCTCAATACACAAGTCTTTATGGACCAACTGATGAGACTCTGTGA | 67 | Db | 1382 | TAAATGGTGTATGCTTAATGGGTCCATACCAACATCAGAGCCTTTGAAATATATCGCAAAATGT | 1441 |
| QY | 958 | GCTTAGGAGATACCGATCTTTGGCGAGAGTAGAATCACTATACCACTATGGCGAAG | 1017 | QY | 2329 | TTGGGATCACGGCAAGCGGAAATTTGACACCAAGATCACTTTGTTTCCAAAGTCGCT | 2388 |
| Db | 68 | GATTAGGAGATACGAATCTTTTGGACAGATTCGGAATGAAAAGACTATGCAATTTATGGAGATG | 127 | Db | 1442 | ATGGTCAATTTGGTGGTAAACATTACACATATGCTGTATGACTTTGTTTCTAACACATGCAAT | 1501 |
| QY | 1018 | AACCTAAATTTGGCGGGTAAACATATCCGTAGAGGTATGGGTGAGA-----GCA | 1068 | QY | 2389 | ATGAAAATGGCGTGAAGAAAAGCTGGGCTTAGAGCGCAAGTTCTACCGGTCAAAAACCT | 2448 |
| Db | 128 | AAGCTACTTTTCGTGGCGGAAATCAATTCGTGATGGTCTCAAAATCCCTAAATGTGA | 187 | Db | 1502 | ATGAAAACGGTATTTATTCGTCACCTCAATCTTAAACGAATGTTTCGACAGTTAGAAATA | 1561 |
| QY | 1069 | ATAGCCCTCATGAAACACCTAGATTTAGTCACTCACTAACGCGATGATTTACGACTACA | 1128 | QY | 2449 | GCCGTAAATCACCAAGAAAGACTTCAAGTTTCAACGACAAACCGCAAAAAATCACCGTCG | 2508 |
| Db | 188 | CAAGAGATGATAAAAAATGTAGCCGATTTAGTTTAACTAACGCAATTAATTTATTTATG | 247 | Db | 1562 | TTAGAAATTTAACTAAGCGCAGATATGAAAAATAATAATGCTACACCTTAAATATAGATGTAG | 1621 |
| QY | 1129 | CGGGATTTTACAAACCGGACATTTGGGATTAATAAACGCAAAATCCATGGCATTTGGCAAG | 1188 | QY | 2509 | ATCCGAAAAACCTTCGAGGTCTTTGTAGATGCAAAACTCTGCACCTCTTAAACCCACCTCGC | 2568 |
| Db | 248 | ACAGATTTGTTAAGCAGATATCCGAAATTAATAATGGTTATATTTTAAAGATCGTTAAG | 307 | Db | 1622 | ATCCACAAACATATGAGGTATTCGTTGATGTTAAATAATAATCACAGTGAAGCAGCAACAG | 1681 |
| QY | 1189 | CAGGAAACCAAGGACATGCAAGATGGCGTAAAGCCCTCATATGTCGTGGGTGTTGGGCACAG | 1248 | QY | 2569 | AAGTGCCTCTAGCCCGCGCTACACTTTTCTCTAG 2603 | |
| Db | 308 | CTGGAACCCAGATATAATGATAACGT-----TGACATCATCAATTGGTGCAACACTG | 361 | | | | |
| QY | 1249 | AAGCACTAGCAGGGAAGGTATGATTTATACCGCTGGGGAATCGATTTCACACCCCACT | 1308 | | | | |
| Db | 362 | ATATTATTGCTGTGAAGGTAAATTTGTTACTGCCGCGGTATCGATACACACGTCGCACT | 421 | | | | |
| QY | 1309 | TCCTTTCTCCACAAATTCCTTACCGCTCTAGCCATGCGGTTTACCAACCATGTTTGGAG | 1368 | | | | |
| Db | 422 | TCATCAATCTTGAAACAGCTGAAGTTGCACTTGAGAGTGGTATTTACACCGCATATCGGTG | 481 | | | | |
| QY | 1369 | GCGGCAAGGCTCTGTAGATGGCAGCAATGCGACTACTATCACTCCGGGCAAAATGGAACCT | 1428 | | | | |
| Db | 482 | GAGGAATGTTCTTCTGAAGTGTCTTAAGCGACTACTGTAAACACGAGACCTTGGCATA | 541 | | | | |
| QY | 1429 | TGCACCGCATTTTGGCGGACGAGAAAGATTTCTATGAATTTGGGCTTTTGGGCAAG | 1488 | | | | |
| Db | 542 | TTTCATGCGCATTTTGAAGACGAGAGAGATGCTTAAATGATGAGGATTTACTGGTAAAG | 601 | | | | |

Db 1682 AATTACCATTAAACACAAAGATACCTCTTATTCTAG 1716

RESULT 4

US-10-485-517-26
; Sequence 26, Application US/10485517
; Publication No. US20050256299A1
; GENERAL INFORMATION:
; APPLICANT: University of Sheffield
; APPLICANT: Biosynexus Incorporated
; APPLICANT: Foster, Simon
; APPLICANT: Mond, James
; TITLE OF INVENTION: Antigenic Polypeptides
; FILE REFERENCE: P100629W0
; CURRENT APPLICATION NUMBER: US/10/485,517
; PRIOR FILING DATE: 2004-02-02
; PRIOR APPLICATION NUMBER: GB 0118825.9
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: GB 0200349.9
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 424
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26
; LENGTH: 1612
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-485-517-26

Query Match 15.8%; Score 455; DB 1; Length 1612;
Best Local Similarity 57.5%; Pred. No. 4.1e-101;
Matches 928; Conservative 0; Mismatches 630; Indels 55; Gaps 4;

QY 619 ATCTTGCATGCGGTAGCAATTCACATTCCTTTGAAGCTTAACAGGCATTAATAATTGCA 678
Db 6 ACCAATTCAAGTGGGCTCAATTTTCATTTTATGAAGCAATTCAGCATTTAGATTTCGA 65
QY 679 TCGTGAAGAACCTATGGCAACCGCTAGATATTTCCCTCTGGCAACACGCTAGCATTTGG 738
Db 66 ACGTGAATGGCATATGGAAACATTTTAGATATTCAGCTGGAGCAGCTGTCGATTGA 125
QY 739 GGAGGACAAACCGCAAGTCAGATTGATTCCTCTTGTGGCAGTAAATAAGTATTGG 798
Db 126 ACTGGGGATAAAGAGAGTTCAATTAGTTGAATATGCTGGCAACAGCTAAATTTTGG 185
QY 799 CATGAACGGCTTTGTGAA---TAAACATCGGGATGAACGCCATAAACATAAGCGTTGA 855
Db 186 TTTTCGTGGTATGGTCAAATGGTCTATCGATGATCAGCTGTCTATCGCCCAACTGATGA 245
QY 856 CAAGGCGAATCTCAGGATTATCA----- 881
Db 246 AATGATGAATATGCGAGGTGATTCGGAGATAACGGTGTGAACCGTGAATAAAAAGG 305
QY 882 ---AGTAAGGAGACTCCCATGAATAATGAATAAACAAGAAATATGTAATACCTACGGACCC 938
Db 306 AGGAAAAGATCATGAGCTTTAAATATGAACCAAAATCAATATACGAGCTTATACGGTCCA 365
QY 939 ACCAAAGGCGAATAAGTCGCTTTAGGAGATACCGATCTTTGGCAGAGTAGAACATGAC 998
Db 366 ACTGTTGGAGATTCATTCGTTTAGGTGATACGAAATCTATTTGCTCAAAATAGAAAAGAC 425
QY 999 TATACACCTATGGCAGAACCTAAATTTGGCGGGGTAAACTATCCGTGAGGGGTATG 1058
Db 426 TATCGCGTTTATGGTGAAGAGCTACTTTTGGTGGTAAATCTATTAGAGACGGTATG 485
QY 1059 GGTACAGCAATAGCCC-----TGATGAAAACACCCCTAGATTTTGTGATCATCACTAAC 1109
Db 486 GCGAAAATCCTCGTGTAAACACTGATGACGTGAACGTTCCAGACCTTGTTCATTTCTAAT 545
QY 1110 GCGATGATTATCGACTACACCGGGATTTTACAAAGCCGACATTTGGGATTTAAAAACGGCAAA 1169
Db 546 GCGGTTATTATCGATTACGATAAAGTGGTTTAAAGCTGATATAGCATTAATAAATGGTTAT 605

RESULT 5

US-10-793-626-4032
; Sequence 4032, Application US/10793626
; Publication No. US20050255478A1

QY 1170 ATCCATGGCATTGGCAGGAGGAAACAAGACATGCAAGATGGCGTAAGCCCTCATATG 1229
Db 606 ATTTTCGCCATAGGTAATGCGCGCAACCCAGCATTAATTAATGT-----CGACATT 659
QY 1230 GTCGTGGGTGGGCACAGAAAGCACTAGCAGGGAAGGTATGATTATTACCGCTCGGGGA 1289
Db 660 ATTATAGGTTCAACAACAGATATCATTCGCCCTGAAGGTAAATCGTCACTGCTGGTGT 719
QY 1290 ATCGATTACACACCCCACTTCCTTTCTCCAAACAATTCCTTACCGCTCTTAGCCAAATGCG 1349
Db 720 ATTGATACTCATGTTTCATTTTATTAACTCTGAACAGCAGAGGTGCGATTAGAAAGTGT 779
QY 1350 GTTACAAACCATGTTTGGAGCGGCACAGGTCTCTGTAGATGGCAAAATGCGACTACTATC 1409
Db 780 ATTACGACTCATATTTGGTGGTGTACTGGTCTTCAGAAAGGTTCTTAAAGCAACAACGTGA 839
QY 1410 ACTCCGGGCAATGGAACCTTCACCGCATGTTGCGCGCAGCAGAGAGTATTCTATGAAT 1469
Db 840 ACTCCAGGTCCATGGCATATTCATAGAATGTTAGAGCTGCCGAAGGTTTACCGATTAAT 899
QY 1470 GTGGGCTTTTGGGCAAAAGGCAATAGCTCTTAGCAAAAAACAATCTGTAGAAACAAGTAGAA 1529
Db 900 GTCGTTTTACAGTAAAGGACAAGCAACAATCCAATGCACTCATTTGAACAATCAAT 959
QY 1530 GCGGCGCGATTGTTTTAAATTCATGAAGACTGGGGCAACAACCAAGTGCATCGAT 1589
Db 960 GCGGAGCAATTGGATTAAAGATGATGAAGACTGGGGTGCAACACCATCTGCTTTGAGT 1019
QY 1590 CACTGCTTGGCGTGGCAGATGAATACGATGTCGAAGTTTGTATCCACACCGATCAGT 1649
Db 1020 CATGCATTAGATGTTGCTGATGAATTTGATGTTCAAAATTCATATCATGACAGATACTTTA 1079
QY 1650 AATGAGCAGGTTATGATGACACCTTAATGCAATGAACGCGCGCCATCCATGCC 1709
Db 1080 AATGAAGCAGATTATGGAAGACACAATGCTGCTGTTAAAGCCGTGTACTTCATATG 1139
QY 1710 TACCACATTGAGGAGCGGGTGGAGGACACTCACCTGATGTTATCACCATGSCAGCGAG 1769
Db 1140 TACCATATGAGGTGCTGGTGGGTCATGGCTGATTTAAATTTAAATTCGCTGATTT 1199
QY 1770 CTCAATATTCCTCTCCACACCCCACTATTCCTCTATACCAATTAACGGTTGCA 1829
Db 1200 TCAAAATATTTACCTTCATCTACAAATCCAACTTTGCCCTTATACACATAACTCTGTAG 1259
QY 1830 GAACACTTAGACATGCTCATGACATGCGCACCACTAGACAAACGATCCGCGAGGATTTA 1889
Db 1260 GAACATTTAGATGTTGTAATGATTACTCACCATTAAATGCGGCTATTTCCTGAAATATC 1319
QY 1890 CAATTTTCTCAAAGCCGTATCCGCGCGGCTCTATCGCGGCTGAAGATGTGCTCCATGAT 1949
Db 1320 GCATTCGAGATTCACTGTTTCGTAAGAAACGATTGCGACGAGAAGATTCTGCAAGAT 1379
QY 1950 ATGGGTGTATCGCGATGAACAAGTCGGATTTCGCAAGCAATGGGCGGTGCGAGCGAAGTG 2009
Db 1380 ATGGGTGTATTCAGTATGATTAGTTCCGATTTCACAAAGCAATGGGCGGTGAGGTGAAGTA 1439
QY 2010 ATTCTCTGAACCTTGGCAGACTGCGGATGAAGATTAATAAAGAAATTTGGTAAGCTTCTGAA 2069
Db 1440 ATTAACGAAACATGGCAAGTAGCACATCGCATGAAAGAAACAACGCGGTCTTTAGATGGT 1499
QY 2070 GATGGCAAGATACGATAATTTCCGCATTAAGCGCTACATCTCCAAATACACTATCAAC 2129
Db 1500 GATTTGAACATAATGATAATTAATCGCATCAAAAGTTATATCGCTTAATATACATTAAC 1559
QY 2130 CCGCTTTTGACCCACCGCGGTGAGCGAGTATATCGGCTCTGTGGAAAGAGGGCAA 2182
Db 1560 CCAGCAATTACACATGGTATTCTGTAATATGTAGGATCTATCGAGCCGGGCAA 1612

GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4032
; LENGTH: 3475
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-4032

Query Match 14.7%; Score 424.4; DB 1; Length 3475;
Best Local Similarity 55.1%; Pred. No. 1.6e-93;
Matches 1010; Conservative 1; Mismatches 762; Indels 59; Gaps 7;

QY 7 ATTTCARCACTCAAGCACATATTGATCCTGTGTGTGGTGGTAAATTCRACCTGT 66
Db 1653 AATATTTCATCATCAATTACATTTTACTAAATAATGTTTATGTACGCTGGATATTTT 1712

QY 67 TAAATCTATTATTAATTTTTTAATAATTAATTAATTAATTAATTAATTAATTAAT 126
Db 1713 CATATTATTCAACTTATACACTTAATTTTAAATAAACATTTGTAATTAATTAATTAAT 1772

QY 127 ATATTAATAAAGTTAATAAAGTAACGAAATAGGACTATAATCCCATTCGCTTTAAAT 186
Db 1773 CAAGTATTATCTGTATCTTCAATTAATTAATTAATTAATTAATTAATTAATTAAT 1932

QY 187 TTAACACAAAGGA---GTAATAGGTGAACTCACACCAAGAGCAAGAAAGTTCTTGTT 243
Db 1833 TAATATGAAGAAGGATAGTTTGCATTTTACACAACTGTAACAGCAAAATTTGATGAT 1892

QY 244 ATATTAATGCGGGAAGTGGCTAGAAAGCGCAAGAGAGGGCTTAAAGCTCAACCAAC 303
Db 1893 AGTTGTAGCTGCTGAGGTTGCACGCTGTAGAAAAGCAAGAGGACTTTAAACTTAATCATCC 1952

QY 304 CGAAGCATTGCTTACATTAGTCCCATATTATGAGCAAGCGCGCTGGAAAATAAAC 363
Db 1953 TGAAGCATTGCTTTAATCAGTATGAATTAATGAGGCGCGCTGATGG---TAAAC 2009

QY 364 CGTTGCCAGCTTATGGAAGAGTGCATGCACCTTTTGAATAAAGATGAAGTAATGCCCG 423
Db 2010 GGTAGCTGAATCATGAGCTATGGAATAACAAATTTTAAACGAGAGATGTCATGGATGG 2069

QY 424 GGTGGTAATATGTTCCCGATCTAGGTGTAGAGCCACCTTTCTTCATGGTACGAACT 483
Db 2070 CGTAGCTAATCATGATTACAGAACTTGAAATTTGAAGCAACTTTTCCAGATGGTAAAGTT 2129

QY 484 TGTAACTGTGAATGCGCCCATCAACAGATGAGCACTTCAAGCGGGC---GAAGTGAAT 542
Db 2130 AATAACAGTCCCATCACCCATCTTTAAGAGGCTATATCAATGATTCCTGGTGAATTA 2189

QY 543 TTGGTTGGCATTAAGACATCGAGCTCAATGTCAGGCAAGAAAGTAACCGAACTTTGAGGTTA 602
Db 2190 TTGTTAAATAACTGAATAGAGTCAATAAACATCATCCGAAACGGTGAATGAAGTAA 2249

QY 603 CTAATGAAGGCCATAATCTTCATGTGGGTAGCCATTTCCACTTCTTTGAAGCTAAACA 662
Db 2250 AAAATACAGCGATAGACCTATACAAAGTAGTGTTCACATTTCCACTTTTTCGAAGCAATA 2309

QY 663 AGGCATTAATTCGATCGTGAAGAAAGCCATATGCGCAACCCCTAGATATTCCTCTGGCA 722
Db 2310 AAGCATTAGAATTTGATCGTGAAGAAAGCATATGGTAAACATTTGGATATTTCTCGCAGGAG 2369

QY 723 ACAGCTACGATTTGGGGCAGGACAAACCCGCAAGATGCAAGTTGATTCCTCTTGGTGGCA 782

Db 2370 CTGCACTGAGATTGAACTGGAGATGAAAAAAGATACAACTTTGTGCAATATTCTTGGAC 2429

QY 783 GTAAAAAAGTATTGGCATGAACGGCTGTGTAAT---AACATCGCGATGAACGCCATA 839

Db 2430 GACGTAAATTTATAGGATTCGGTGGTTTGTAGTCGATGGGATATTGACGAAGAACGGGTAT 2489

QY 840 AACATAAAGCGCTTGAACAAGCGGCAAAATCTCACGGATTTATCAAGTAAGGAGACTCCCATG 899

Db 2490 TCCGTCCAAATGATTCAAAATCAAAAACGCCCGCTTAAAAACGATGCAGCGGAAGACAATG 2549

QY 900 -----AAAATGAAAAAACAAGATATGTAAA 925

Db 2550 CGAATAAAAAAGGTGTAATAATAGCATGAGTTTAAAAATGACAAATCTCAATACACAAG 2609

QY 926 TACCTACGGACCCACCAAGCGGATAAAGTCGCTTAGGAGATACCGATCTTTGGGCGAGA 985

Db 2610 TCTTTATGGACCAACTGTAGGAGACTCTGTGAGATTAGGAGATACGAACTTTGTTTGCACA 2669

QY 986 AGTAGAACATGACTATACCACTTATGCGGAAGAACTTAAATTTTGGCGCGGTAAAACTTAT 1045

Db 2670 AGTTGAAAAAGACTATGCAAAATTTATGAGATGAAGCTACTTTTCGGTGGCGGAAATCAAT 2729

QY 1046 CCGTAGGGTATGGGTGAGA-----GCAATAGCCCTGATGAAAAACACCCCTAGATT 1096

Db 2730 TCGTGTGTTGTTGGCTCAAAATCCTAATGTGCAAGAGATGATAAAATTTAGCCGCTTT 2789

QY 1097 AGTCATCACATAACCGGATGATTATCGACTACACCGGATTTTACAAAGCGCATTTGGGAT 1156

Db 2790 AGTTTAACTAAACGATTAATTTATGATTTATGACAAGATTGTTAAAGCAGATATCGGAAT 2849

QY 1157 TAAAAACGGCAAAATCCATGGCATTTGCAAGCGGAGGAAACAAAGGACATGCAAGATGGGAT 1216

Db 2850 TAAAAATGGTTATATTTTAAAGATCGGTAAAGCTGGAACCCAGATATAATGATAACGT 2909

QY 1217 AAGCCCTCATATGGTGTGGGTGGGCAAGAAAGCACTAGCAGGGGAAGGTATGATTAAT 1276

Db 2910 -----TGACATCATCTATGTTGCAACAACATGATTAATTTGCTGCTGAAGGTAAATTTGT 2963

QY 1277 TACCGTGGGGGATCGATTTCACACACCCACTTCTTCTCCACAACAATTTCCCTACCGC 1336

Db 2964 TACTGCGCGGGTATTCGATACACACGTCGACTTTCATCAATCCTGAACAGCTGAAGTTGC 3023

QY 1337 TCTAGCCAAATGGCGTTTACAAACCATGTTTGGAGGCGGCACAGCTCCTGTAGATGGCAGAA 1396

Db 3024 ACTTGAGAGTGTATTACAAAGCATATCGGTGGAGGAACTGCTGCTTCTGAAGGTGCTAA 3083

QY 1397 TGGCACTACTATCACTCCGGGCAAAATGGAACCTTGCACCGCATGTTGCGCGCAGCAGAGA 1456

Db 3084 AGCGACTACTGTAAACACACGAGCCCTTGGCATATTTTCATCGCATGTTAGAAGCAGCAGAGA 3143

QY 1457 GTATTCTATGATGTGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAACACTTGT 1516

Db 3144 GATGCTCTTAATGATAGGATTTTACTGTTAAAGGTCAAGCTGTCAATCATATCTGCACTTAT 3203

QY 1517 AGAACAGTGAAGCGGCGGATTTGTTTAAATTTGATGAAGACTGGGGCACAACACC 1576

Db 3204 TGAACAAATTCATGCGAGCGCTATAGGTCTTTAAAGTACATGAGATTGGGAGCTACACC 3263

QY 1577 AAGTCGATCGATCACTGCTTTAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCA 1636

Db 3264 TTCAGCATTAAGTCATGCAATTAGACGTTGTCAGATGAGTTTCAAGTCGCTTTACA 3323

QY 1637 CACCGATACAGTCAATGAGGCAAGTTATGATAGTACACCCCTAAATGCAATGAACGGGCG 1696

Db 3324 TGCAGATACATTAATGAAAGCTGGAATTTATGGAAGATACAAATGGCTGCTGTGTAAGATCG 3383

QY 1697 CGCCATCCATCCCTACCACTTGAAGGAGCGGGTGGAGGACACTCACCTGATGTTTATCAC 1756

Db 3384 TGTATTGATATGATCATCTGAAGGAGCTGGGTGGTGGTATGATGATGATGATGATGAT 1757

QY 1757 CATGGCAGGCGAGCTCAATATTCTACCCCTCCT 1788

Db 3444 ATCAGCTGCATATTCAAAACATCTTACCTTCT 3475


```
QY 774 TTGTGGCAGTAAAGATGTTGGCATGAACGGCGTTGTGAATACATCGCGGATGAA 832
Db 185 ATTCTGGACGAGTAAATTTATGAGTTCGCGTGGTTAGTCGATCGGATATTGACGAA 243

RESULT 9
US-10-793-626-4078
; Sequence 4078, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4078
; LENGTH: 2259
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-4078

Query Match 1.6%; Score 46; DB 1; Length 2259;
Best Local Similarity 48.2%; Pred. No. 0.061;
Matches 158; Conservative 1; Mismatches 166; Indels 3; Gaps 1;

QY 7 ATTTTCCARCACTTCAAGCACATATTGATCCCTGTGTGTGGTGGTAAATTRCRACCTTGT 66
Db 1907 AATATTCATCATATCAATTAATTTACTAAATAATGTTAATGACGTGGATATTTT 1966

QY 67 TAATRCATATTAAATTTTTTAATAATTACTTTATATCATATATAATAATATTACTT 126
Db 1967 CATATTATTCAACTTTATACACTTAATTTAAAAAACAATGTTAATTAGATAAAATAT 2026

QY 127 ATATTAAAGTTAATAAAAGTAACGAATTAAGGACTATATCCCAATTCCTTTAAAT 186
Db 2027 CAAAGTATTATATCTTGTTACTTCAATTAATAATAGTTTAAATAAAATTTGCAAGACAAAT 2086

QY 187 TTAACACAGGA--GTAATAGTGAACTCACACCAAGAGCAAGAAAAGTCTTGT 243
Db 2087 TAATATGAAGAAGGATAGTTTTCACCTTTACACAACTGTAACAGACAAATTTGATGAT 2146

QY 244 ATATTATGCGGCGAAGTGGCTAGAAAGCGCAAGAGGGCTTAAAGCTCAACCAACC 303
Db 2147 AGTTGTAGCTGCTGAGTTGACGCTGTAGAAAAGCAAGGACCTTAACCTTAATCATCC 2206

QY 304 CAAAGCATTGCTTACATTAGTCCCAT 331
Db 2207 TGAAGCACTTGCTTTAATCAGTGATGAT 2234

RESULT 10
US-10-793-626-4389/c
; Sequence 4389, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4389
; LENGTH: 3373
; TYPE: DNA
```

```
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-4389

Query Match 1.4%; Score 40; DB 1; Length 3373;
Best Local Similarity 56.8%; Pred. No. 2.1;
Matches 67; Conservative 3; Mismatches 48; Indels 0; Gaps 0;

QY 3 GRAGATTTCCARCACTTCAAGCACATATTGATCCTGTGTGTGGTGGTAAATTRCRAC 62
Db 2864 GAGGATTATACACGTTTGTGAAAAATATATTAAATTTAAACAGTGATTGCTAAATCACAAC 2805

QY 63 TTGTTAATRCATATTATTAAATTTTTTAATAATTACTTATTATCATATATAATAATTATTA 120
Db 2804 TTGTCATTATAAATTTATTTTTGTACTATTACAATAATGATTATTGACGTTAATTTA 2747

RESULT 11
US-10-793-626-3926/c
; Sequence 3926, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 3926
; LENGTH: 3927
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3926

Query Match 1.4%; Score 40; DB 1; Length 3927;
Best Local Similarity 56.8%; Pred. No. 2.3;
Matches 67; Conservative 3; Mismatches 48; Indels 0; Gaps 0;

QY 3 GRAGATTTCCARCACTTCAAGCACATATTGATCCTGTGTGTGGTGGTAAATTRCRAC 62
Db 1836 GAGGATTATACACGTTTGTGAAAAATATATTAAATTTAAACAGTGATTGCTAAATCACAAC 1777

QY 63 TTGTTAATRCATATTATTAAATTTTTTAATAATTACTTATTATCATATATAATAATTATTA 120
Db 1776 TTGTCATTATAAATTTATTTTTGTACTATTACAATAATGATTATTGACGTTAATTTA 1719

RESULT 12
US-10-793-626-4109/c
; Sequence 4109, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4109
; LENGTH: 3923
; TYPE: DNA
```

```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-4109

Query Match      1.4%; Score 39.8; DB 1; Length 3923;
Best Local Similarity 55.6%; Pred. No. 2.6;
Matches 74; Conservative 1; Mismatches 58; Indels 0; Gaps 0;

QY   53 AAATTCRACCTGGTAAATRCATTATTATAATTTTAAATTAATTAATCATATATAA 112
    ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB   2103 AATATGATCACTTTTAGTATATGACTATTATTTTATATAAATTAACACTCCTATTAA 2044

QY   113 TAAATATTACTTATATATAAAAAAGTTTAATAAAAAGTAACGAATTAGGACTATAATCCC 172
    ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB   2043 CTGTACTACAGTCAAGATAAGAGTGCGAATAAATAGTACTTTGAATAATTACTATAATTAC 1984

QY   173 ATTGCCCTTTAAAA 185
    | |||| 
DB   1983 ACAAGTATTATAA 1971
    |||| 

RESULT 13
US-10-793-626-2361
; Sequence 2361, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2361
; LENGTH: 1809
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-2361

Query Match      1.3%; Score 37.8; DB 1; Length 1809;
Best Local Similarity 53.3%; Pred. No. 5.3;
Matches 72; Conservative 3; Mismatches 60; Indels 0; Gaps 0;

QY   23 AGCACATATTGATCCCTGTTGTGGGTGGTAAATTRCRACCTTGTTAATRCTATTATTAAT 82
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB   322 AGCTGTCATGGAGTTTTTTTAGTACCTGAAATTTTACAATGAGTACTACTATTCAA 381
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY   83 TTTTAAATAATTACTTATTATCATATATAATAATTATTACTTATATATAAAGTTTAAT 142
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB   382 TCITTTTATCGATTCTAATGATGATTTAAACGATATCGAATTCGATTTGAAATGATTAAAT 441
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY   143 AAAAAGTAACGAAT 157
    |||| 
DB   442 GAGAAACGTCACAT 456
    |||| 

RESULT 14
US-10-793-626-4033/c
; Sequence 4033, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2005, 00:04:09 ; Search time 522 Seconds
(without alignments)
9817.453 Million cell updates/sec

Title: US-09-904-994B-1

Perfect score: 2883
Sequence: 1 rgragatttccarctt.....aaaaagtagaccacagg 2883

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/prodata/1/ina/1 COMB.seq.*
2: /cgn2_6/prodata/1/ina/5 COMB.seq.*
3: /cgn2_6/prodata/1/ina/6A COMB.seq.*
4: /cgn2_6/prodata/1/ina/6B COMB.seq.*
5: /cgn2_6/prodata/1/ina/H COMB.seq.*
6: /cgn2_6/prodata/1/ina/PCTUS COMB.seq.*
7: /cgn2_6/prodata/1/ina/PP COMB.seq.*
8: /cgn2_6/prodata/1/ina/RE COMB.seq.*
9: /cgn2_6/prodata/1/ina/backfileseq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|---------|----|---------------------|
| 1 | 1075.2 | 37.3 | 2735 | 2 | US-08-920-095-1 |
| 2 | 1075.2 | 37.3 | 2735 | 6 | PCT-US96-05800-1 |
| 3 | 1062.6 | 36.9 | 2619 | 2 | US-08-467-822-19 |
| 4 | 1062.6 | 36.9 | 2619 | 3 | US-08-432-697-19 |
| 5 | 1062.6 | 36.9 | 2619 | 3 | US-08-466-248-19 |
| 6 | 1058.6 | 36.7 | 4824 | 3 | US-09-431-705-1 |
| 7 | 1058.6 | 36.7 | 4824 | 3 | US-09-431-705-19 |
| 8 | 737.6 | 25.6 | 6131 | 2 | US-07-732-242C-8 |
| 9 | 693.4 | 24.1 | 2400 | 2 | US-08-967-513-1 |
| 10 | 693.4 | 24.1 | 2400 | 2 | US-08-687-645B-1 |
| 11 | 664.4 | 23.0 | 1830121 | 3 | US-09-557-884-1 |
| 12 | 664.4 | 23.0 | 1830121 | 3 | US-09-643-990A-1 |
| 13 | 664.4 | 23.0 | 1830121 | 3 | US-10-158-863-1 |
| 14 | 662.8 | 23.0 | 5966 | 3 | US-08-956-171B-22 |
| 15 | 662.8 | 23.0 | 5966 | 3 | US-08-781-986A-22 |
| 16 | 653 | 22.7 | 1710 | 3 | US-09-543-681A-1857 |
| 17 | 630.6 | 21.9 | 8729 | 3 | US-09-453-702B-258 |
| 18 | 630.6 | 21.9 | 8729 | 3 | US-10-114-170-258 |
| 19 | 630.6 | 21.9 | 87563 | 3 | US-09-453-702B-57 |
| 20 | 630.6 | 21.9 | 87563 | 3 | US-10-114-170-57 |
| 21 | 623.4 | 21.6 | 1716 | 3 | US-08-487-429A-3 |
| 22 | 623.4 | 21.6 | 1716 | 6 | PCT-US96-05320A-541 |
| 23 | 617.8 | 21.4 | 1878 | 3 | US-09-489-039A-2045 |
| 24 | 612 | 21.2 | 3164 | 3 | US-09-710-279-3828 |

ALIGNMENTS

RESULT 1

US-08-920-095-1
; Sequence 1, Application US/08920095
; Patent No. 5837240
; GENERAL INFORMATION:
; APPLICANT: Cynthia K. Lee et al.
; TITLE OF INVENTION: MULTIMERIC, RECOMBINANT UREASE VACCINE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/920,095
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/431,041
; FILING DATE: 28-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06132/013001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2735 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
US-08-920-095-1

Query Match 37.3%; Score 1075.2; DB 2; Length 2735;
Best Local Similarity 66.0%; Pred. No. 5.8e-267;
Matches 1622; Conservative 0; Mismatches 798; Indels 38; Gaps 3;

Sequence 3853, Ap
Sequence 117, App
Sequence 2189, Ap
Sequence 1786, Ap
Sequence 10316, A
Sequence 10546, A
Sequence 13, Appl
Sequence 15, Appl
Sequence 89, Appl
Sequence 2, Appli
Sequence 1, Appli
Sequence 4032, Ap
Sequence 10513, A
Sequence 10653, A
Sequence 153, App
Sequence 159, App
Sequence 7520, Ap
Sequence 1913, Ap
Sequence 1761, Ap
Sequence 543, App
Sequence 121, App

QY 182 AAAATTTAACACAGGAGTAATAGTGAATCAACCCAAAGAGCAAGAAAGTTCCTG 241
Db 78 ACACCTTTAAGATATAGGAGAACTGAGATGAATCAACCCAAAGAGTTAGATAAGTTGATG 137
QY 242 TTATATTATGCGGCGAAGTGTCTAGAAAGCGCAAGCAGAGCGCTTAAAGCTCAACCAA 301
Db 138 CTCACACTACGCTGGAGAAATTTGGCTAAAACCGAAGAAAGAGCAATTAAGCTTTAACTAT 197
QY 302 CCCGAAGCCATTGCTTACATTAGTCCCATATATTAGGCAAGCGCGCGTGGAAAAAAA 361
Db 198 GTAGAAGCAGTAGCTTTGATTAGTCCCATATATTAGGAAGAGCGAGAGCTGGTAAAAAG 257
QY 362 ACCGTTGCCAGCTTATGGAAGAGTGATGCACTTTTGGAAAAAGATGAATGATGCC 421
Db 258 ACTGCGGCTGAAATGATGCAAGAGCGGCACCTCTTTTAAAAACCATGATGTGTGGAT 317
QY 422 GGGGTGGGTAAATATGGTTCCCATCTAGGTGTAGAGCCACCTTTCCTCATGGTACGAA 481
Db 318 GCGGTGGCAAGCATGATCCATGAAGTGGGTAATTGAAGCGATGTTTCTGATGGGACTAAA 377
QY 482 CTTGTAACTGTGAATTTGGCCCATCGAAACCATGATGAGCACTTCAAAGCGGCGCAAGTGAAA 541
Db 378 CTCGTAAACCGTGCAACCCCTATTGAGGCCAATGGTAAATTAGTTCCTGGTGGTTG --- 434
QY 542 TTTGGTTGGATAAAGACATCGAGCTCAATGAGGCAAGAAAGATACCGAACTTTGAGGTT 601
Db 435 TTTCTTAAAAAATGAAGACATCACTATCAACGAAGGCAAAAAAGCCGTTAGCGTGAAGTT 494
QY 602 ACTAATGAAGGCGCTAAATCCTTGATGTGGGTAGGCCATTTCACACTTCTTTGAAAGCTAAC 661
Db 495 AAAAATGTTGGCGCAGACCGGTTCAAATCGGCTCACACTTCATTTCTTTGAAAGTGAAT 554
QY 662 AAGGCACATAAAATTCGATCGTGAAGAAAGCCTATGGCAACCGCTAGATATTCCTCTGGC 721
Db 555 AGATGCTTAGACTTTGACAGAGAAAAAACTTTCGGTAAACGCTTAGACATTGGCAGCGG 614
QY 722 AACACGCTACGATTGGGCGAGACAAACCCGCAAGTCAGTTGATTCCTCTGTGGTGC 781
Db 615 ACAGCGTAAAGATTGAGCTGGCGAAGAAATCCGTAGATTGATTGACATTGGCGGT 674
QY 782 AGTAAAAAGTGAATGGCATGAACGGGCTTGTGAATAACATCGCGATGAACGCCATAAA 841
Db 675 AACAGAAGAATCTTTGGATTAAACGCAITGGTTGTATAGACAAGCAGACAAACGAAGCAA 734
QY 842 CATAAAGCGCTTGACAAGCGAAATCTCACGGATTT - - - - - 877
Db 735 AAAATTTGCTTTACACAGACTAAAGAGCGTGGTTTTCATGGCGCTAAAAGCGATGACAAC 794
QY 878 - - - - - ATCAAGTAAGGAGACTCCCATGAAA - - - ATGAAAAAACAAGATAATGTAAT 926
Db 795 TATGTA AAAACAATTAAGGAGTAAGAAATGAAAAAGATTAGCAGAAAAAGATATGTTCT 854
QY 927 ACCTACGGAACCAACAAAGCGGATAAAGTGGCTTAGGAGATACCGATCTTTGGGCGAGAA 986
Db 855 ATGTATGGTCTTACTACAGCGGATAAAGTGCAGATTGGGCGATACAGACTTGATCGCTGAA 914
QY 987 GTAGAACATGACTATACCACTATGCGGAGAGCTTAAATTTGGCGGGTAAACATATC 1046
Db 915 GTAGAACATGACTATACCACTTATGCGAAGAGCTTAAATTTGGTGGCGGTAAACCCCTA 974
QY 1047 CGTGAGGGTATGGGTACAGAGCAATAGCCCTGATGAAAAACCCCTAGATTAGTCACTCACT 1106
Db 975 AGAGAAGGCATGAGCCAAATCTAACACCCCTAGCAAGAGAGATTGGAATTAATATCACT 1034
QY 1107 AACCGCATGATTATCGACTACACCGGGAATTAACAAAGCCGACATTTGGGATTA AAAAAGCGG 1166
Db 1035 AACGCTTTAATCGTGGATTACACCGGTAATTAATAAGCGGATATTGGTATTAAGATGGC 1094
QY 1167 AAAATCCATGGCATTTGGCAGGCGAGAAACAAGGACATGCAAGATGGCGTTAAGCCCTCAT 1226
Db 1095 AAAATCGCTGGCATTTGGTAAGCGGTAACAAAGACATGCAAGATGGCGTTAAAAACAAT 1154
QY 1227 ATGGTCTGGGTGTGGGCACAGAAGCACTAGCAGGGGAAGGTATGATTATTATCCGCTGGG 1286

Db 1155 CTTAGCGTAGGTCCTCTACTGAAGCCTTAGCCGGTGAAGGTTTGATCGTAACGGCTGGT 1214
QY 1287 GGAATCGATTACACACCCACTTCCCTTCTCCACAACAATTTCCCTACCGCTCTTAGCCCAAT 1346
Db 1215 GGTATTGACACACATCCACTTCATTTTACCCCCAACAAATCCCTACAGCTTTTGCAGC 1274
QY 1347 GCGGTTACAAACCAATGTTGGAGCGGCACAGGTCCTGTATAGTGGCAGCAATGCGACTACT 1406
Db 1275 GGTGTAAACAACCATGATGGTGGAAACCGGTCCTGCTGATGGCACTTAATGCGACTACT 1334
QY 1407 ATCACTCCGGGCAATGGBRACCTTGCACCGCATGTTGGCGGCAGCAGAGAGTATCTATG 1466
Db 1335 ATCACTCCAGGCAAGAAATTTAAATGATGCTCAGAGCGGTGAAGAAATATCTATG 1394
QY 1467 AATGTGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAAACAACCTTGTAGAACAAAGTA 1526
Db 1395 AATTTAGGTTTCTTGGCTAAAGGTACGCTTCTAACGATGGGAGCTTAGCCGATCAATTT 1454
QY 1527 GAAGCGGCGCGATTGGTTTAAATTGCATGAAGACTGGGGCACAAACCAAGTGCAGATC 1586
Db 1455 GAAGCGGTGCGATTGGCTTTAAAAATTCAGAAAGACTGGGSCACACTCCTCTTGTGCAATC 1514
QY 1587 GATCACTGCTTGAGCGTGGCAGATGAATACGATGSCAAGTTTGTATCCACACCGATACA 1646
Db 1515 AATCATGCGGTTAGATGTTGCGGCAAAATACGATGTGCAAGTCGCTATCCACACAGACACT 1574
QY 1647 GTCAAATGAGCAGGTTATGTAGATGACACCTTAATGCAATGAACGGGCGCGCATCCAT 1706
Db 1575 TTGAATGAAGCCGGTTGTGTAGAGACACTATGGCTGCTATTGTGTGACGACTATGAC 1634
QY 1707 GCCTACCACTTGAAGGAGCGGTTGGAGGACACTCACCTGATGTTATCACCATGGCAGGC 1766
Db 1635 ACITTTCCACACTGAAGCGCTGGCGGGACACGCTCCTGATATTTATTAAGTAGCGGT 1694
QY 1767 GAGCTCAATATTCTACCTCTCCACACCCGCACTATTCCCTATACCATTAATACGGTT 1826
Db 1695 GAACAAACAATTTCTCCGCTTCCACTAAGCCCAACCATCCCTTTCCCGTGAATACAGAA 1754
QY 1827 GCAGAACCTTTAGACATGCTCAGATGTCACACCTAGACAAACCGATCCGCGAGAT 1886
Db 1755 GCAGAGACATGACATGCTTATGTTGTGCCACCACTTTGGATAAAGCATTTAAAGAAAT 1814
QY 1887 TTACAAATTTTCTCAAAGCCGTATCCGCGCGGCTCTATCCGCGCTGAAGATGTGCTCCAT 1946
Db 1815 GTTCAGTTCCGCTGATTCAAGGATCCGCTCAAACCATTTGCGGCTGAAGACACTTTGCT 1874
QY 1947 GATATGGGTGTATGCGGATGACAAAGCTCGGATTGCGAAGCAATGGGCGGTGAGGCGAA 2006
Db 1875 GACATGGGGATTTTCTCAATCAACAGTTCTGACTCTCAAGCGATGGGCGGTGCGGTGAA 1934
QY 2007 GTGATTCCTCGAATTTGGCAGACTGCGGATAAGATAAAGAAATTTGGTAAAGCTTCCT 2066
Db 1935 GTTATCACTAGAACTTTGGCAACAGCTGACAAAAACAAAGAAAGAAATTTGGCGGCTTGAA 1994
QY 2067 GAAGATGGCAAGATACGATAATTTCCGCAATTAAAGCGCTACATCTCCAAATACACTATC 2126
Db 1995 GAAGAAAAAGCGAATAACGCACTTCAGGATCAACCGCTACTTGTCTAAATACACCAT 2054
QY 2127 AACCCCGCTTTGACCCCAACCGCGTAGGAGTATATCGCTCTGTGGAAAGAGGCGAAGATC 2186
Db 2055 AACCCAGCGATCGCTCATGGGATTTAGCGAGTATGTAGGTTCAGTAGTAAGTGGCAAGTG 2114
QY 2187 GCCGACTTGTGTGTGGAAATCTCGCTTTTGGCGTAAAAACCCAAATCGTGTATCAA 2246
Db 2115 GCTGACTTGTATTGTGGAGTCCAGCATTTCTTTGGCGTGAACCCCAACATGATCATCAAA 2174
QY 2247 GCGGTATGTGTGCTTCTCTGAAATGGCGATTTCTAACCGCTCTGTGCCCACTCCCAA 2306
Db 2175 GCGGATTCATTGGTTAAGCCAAATGGGCGATGCGAACGCTTCTATCCCTACCCACAA 2234
QY 2307 CCGGTTTATTACCGCAAAATGTTTGGGCATCACGCGAAGGCGGAAATTTGACACCGACTC 2366

Db 2235 CCGGTTTATTACAGAGAAATGTTGCTCATCATGTTAAAGCTAAATACGATGCAAAACATC 2294
QY 2367 ACTTTTGTTCCTCAAGTCGCTATGAAATGGCGTGAAGAAAGAAAGCTGGGCTTAGAGCGC 2426
Db 2295 ACTTTTGTGCTCAAGCGGCTTATGACAAAGGCATTAAGAAAGATTAGGACTTGAAGA 2354
QY 2427 CAAGTTCTACCGGTCAAAACTCGCGTAAACATCACCAAGAAAGACTTCAAGTTCAACGAC 2486
Db 2355 CAAGTGTGCGGTAAATAATTCAGAAATATCACTAAATAAGACATGCAATTCACGAC 2414
QY 2487 AAACGGCAAAATCACCGTCGATCCGAAACCTTCAGGTCCTTGTAGATGGCAAACTC 2546
Db 2415 ACTACTGCTCACATGAAGTCAATCTCTGAAACTTACCATGTTGTCGTGATGGCAAGAA 2474
QY 2547 TGCACTCTAAACCCACTCGCAAGTCCTCTAGCCAGCGCTACACTTCTTCTAGG 2604
Db 2475 GTAACTTCTAAACAGCAATAAAGTGAGCTTGGCGCAACTCTTAGCAATTTCTAGG 2532

RESULT 2

PCT-US96-05800-1
; Sequence 1, Application PC/TUS9605800
; GENERAL INFORMATION:
; APPLICANT: OraVax, Inc.
; TITLE OF INVENTION: MULTIMERIC, RECOMBINANT UREASE VACCINE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/05800
; FILING DATE: 23-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/431,041
; FILING DATE: 28-APR-1995
; CLASSIFICATION:
; APPLICATION NUMBER: 08/569,122
; FILING DATE: 06-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06132/020001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2735 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
PCT-US96-05800-1

Query Match 37.3%; Score 1075.2; DB 6; Length 2735;
Best Local Similarity 66.0%; Pred. No. 5.8e-267;
Matches 1622; Conservative 0; Mismatches 798; Indels 38; Gaps 3;
QY 182 AAAATTTAACACAGGAGTAAATAGGTGAAACTCACACCCCAAGAGAGCAAGAAAGTCTCTTG 241
Db 78 ACATTTAGATAGGAGATGAGATGAAACTCACCCCAAGAGATGATGATGATGATGATGATG 137

QY 242 TTATATTATGCGGCGGCAAGTGGCTAGAAAGCCCAAGACGAGGGCTTAAAGCTCAACCAA 301
Db 138 CTCCACTACGCTGGAGAATTGGCTAAAAACGCAAGAAAAAGGCATTAAGCTTAACTAT 197
QY 302 CCGAGGCAATGCTTACATTAGTCCCATATATATGACGAAGCGCGCTGCAAAAAA 361
Db 198 GTAGAAGCAGTAGCTTTGATTAGTCCCATATATATGGAAGAACGAGAGCTGCTAAAAAG 257
QY 362 ACCGTTGCCAGCTTATGGAAGAGTGCATGCTTTTGTGAAAAAAGATGAAGTAATGCCCC 421
Db 258 ACTGCGGCTGAATGATGCAAGAGGCGCACTCTTTTAAACCAGATGATGATGATGAT 317
QY 422 GGGGTGGGTAATATAGTTCCCGATCTAGGTGTAGAACCCACTTCTTCTGATGTAAGAAA 481
Db 318 GCGGTGGCAAGCATGATCCATGAAGTGGGTATTTGAAGCGATGTTTCTGATGGCACTAAA 377
QY 482 CTGTGAACTGTAATTTGGCCCATCGAACGAGATGAGCACTTCAAGCGGCGGAGTGAAGA 541
Db 378 CTGCTAACCGTGCATACCCCTATTGAGGCCAATGGTAAATTAGTTCTCTGGTGAAGTTG -- 434
QY 542 TTTGTTGCGATAAAGACATCGAGCTCAATGCGGCAAGAAAGTAACCCGAACCTTGAGGTT 601
Db 435 TTCTTAAAAAATGAAGACATCACTATCAAGAGGCAAAAAAGCCGTTAGCGTGAAAGTT 494
QY 602 ACTAATGAAGGCGCTTAAATCTTTCATGTTGGGTAGCCATTTCCACTTCTTTGAAGCTAAC 661
Db 495 AAAAATGTTGGCGACAGACCGGTTCAAATCGGCTCACACTTCCATTTCTTTGAAGTGAAT 554
QY 662 AAGGCATAAAAATTCGATCGTGAAGAAAGCTATATGCAAAACGCGCTAGATATTCCTCTGGC 721
Db 555 AGATGCTTAGACTTTGACAGAGAAAAAACTTTTCGGTAAACGCTTAGACATTCGAGCGGG 614
QY 722 AACAGCTACGATTTGGGCGAGGACAAACCCCAAGTGCAGTTGATTTCTCTTGGTGGC 781
Db 615 ACAGCGTAAAGATTGAGCCTGGCGAAGAAAAATCCGTAGAAATGATTGACATTTGGCGGT 674
QY 782 AGTAAAAAAGTATTGGCATGAACCGGCTTGTGAATAACATCGCGGATGAACCCCATAAA 841
Db 675 AACGAGAATCTTTGGATTTAACCCATTTGTTGATAGCAAGCAGACAAACGNAACGCAA 734
QY 842 CATAAAGCGCTTGACAAAGCGGAAATCTCAACGGAATTT----- 877
Db 735 AAAATTGCTTTACACAGAGCTAAAGAGCGTGTGTTTTCATGGCGCTAAAAAGCGATGACAAC 794
QY 878 -----ATCAAGTAAGGAGACTCCCATGAAA-----ATGAAAAACAAAGATATGTAAT 926
Db 795 TATGTAAAAACAATTAAGGAGTAAAGAAATGAAAAAGATTAGCAGAAAAAGATATGTTTCT 854
QY 927 ACCTACGACCCCAAGAGCGATAAAGTGGCTTAGGAGATACCGATCTTTGGGCGAGAA 986
Db 855 ATGATATGTCCTACTACAGGCGATAAAGTGAATTTGGGCGATACAGATTTGATCGCTGAA 914
QY 987 GTAGAAATGACTATACCACCTATTCGCGAAGAACTTAAATTTGGCGGGGTAAAACTATC 1046
Db 915 GTAGAAATGACTACACCATTTATGCGGAGAGCTTAAATTCGTTGGCGGTAAAAACCTTA 974
QY 1047 CGTAGGGTATGGGTGAGAGCAATAGCCCTGATGAAAAACACCCCTAGATTTAGTCACT 1106
Db 975 AGAGAAGGCATGAGCCAACTCTAAACACCTTAGCAAAAGAGATTGGATTTAATTAATCACT 1034
QY 1107 AACGGGATGATTATGACTACACCGGATTTACAAAGCCGACATTTGGGATTTAAAAACGGC 1166
Db 1035 AACGCTTTTAATCGTGGATTTACACCGGTATTTATAAAGCGGATTTGGTATTTAAAGATGGC 1094
QY 1167 AAAATCCATGCTATGGCAAGCAGCAAAAAGAGCATGCAAGATGGCGCTTAAGCCCTCAT 1226
Db 1095 AAAATCGTGGCATTTGGTAAAGCGGTAAACAAAGACATGCAAGATGGCGCTTAAAAACAT 1154
QY 1227 ATGTCGTGGGTGGGCAAGAACACTAGCAGGGGAGGTATGATTAATTAACCGCTGGG 1286
Db 1155 CTAGCGTAGGTCTCTGCTACTGAGCCCTTAGCCGCTGAAGGTTTGTATCTGCTAACCGGCTGGT 1214

QY 1287 GGAATCGATTACACACCCACTTCTTCTTCTCCACAAACAATTCCTACCGCTCTAGCCAAT 1346
Db 1215 GGTATTGACACACATCCACTTCAATTTTCAACCCCAACAATCCCTACAGCTTTTGAAGC 1274
QY 1347 GCGGTTACAACTATTTGGAGCGGCACAGGTCCTGTAGATGGCAGCAATCGACTACT 1406
Db 1275 GGTGTAAACAACCATGATTGGTGGTAAACCGGTCCTGTGATGGCACTAATGCGACTACT 1334
QY 1407 ATCACTCCGGGCAAAATGGAACCTTTGCACCGCATGTTGCGGCGCAGCAGCAAGATTTCTATG 1466
Db 1335 ATCACTCCAGGCAGAGAAATTTAAATGGATGCTCAGAGCGCTGAAGATATTCTATG 1394
QY 1467 AATGTGGGCTTTTGGGCAAGCAATAGCTCTAGCAAAAAACAATTTAGAACAAAGTA 1526
Db 1395 AATTTAGGTTTCTGGCTAAAGGTAACGCTTTTAAACGATGCGAGCTTAGCCGATCAAAAT 1454
QY 1527 GAAGCGGCGCATTTGGTTTAAATGGATGATGAGACTGGGACACACCAAGTGGATC 1586
Db 1455 GAAGCGGCTGCGATTTGGCTTTAAATTCAGAGACTGGGGCACCACTCTTCTTGAATC 1514
QY 1587 GATCACTGCTTGAGCGTGGCAGATGAATACGATGTCAAATTTGTATCCACACCGATACA 1646
Db 1515 ATCATGCGTTAGATGTTGCGGACAAATAGATGCAAGTCGCTATCCACACAGACT 1574
QY 1647 GTCAATGAGCGAGTTATAGATGACACCTTAAATGCAATGAACGGGCGCGCATCCAT 1706
Db 1575 TTGAATGAAGCGGTTGTGTAGAAGACACTATGGCTGCTATTGCTGGACGCACTATGCAC 1634
QY 1707 GCTTACCACATTTAGGGAGCGGTGAGGACACTCACTGATGTTATCACCATGGCAGC 1766
Db 1635 ACTTTTCCACACTGAAGCGCTGGCGGCGACACGCTCTGATATTATTAAGTAGCCGT 1694
QY 1767 GAGCTCAATATTCTACCTCTCCACACCCCACTATTCCCTATACCATTAAATACGGTT 1826
Db 1695 GAACAACACATTTCTCCGCTTCCACTAACCCACCACTCCCTTACCGTGAATACAGAA 1754
QY 1827 GCAGAACACTTAGACATGCTATGATGACACCCACCTAGACAAACGCAATCCGCGAGAT 1886
Db 1755 GCAGAGCACATGGACATGCTTATGTTGTGTCACCACTTGGATAAAAGCAATTAAGAAGAT 1814
QY 1887 TTCAATTTTCAAAAGCGTATCCGCGCGGCTCTATCGGCTGAAAGATGTCCTAT 1946
Db 1815 GTTCAGTTGCTGATTTCAAGGATCCGCTTCAAAACCAATTTGCGCTGAAGACACTTTGCTAT 1874
QY 1947 GATATGGGTGATCGCGATGACAGCTCGATTTCCAGACCAATGGGCGTGCAGCGAA 2006
Db 1875 GACATGGGATTTTCTCAATCACAGTCTGATCTCAAGCGATGGCGGTGGGTGAA 1934
QY 2007 GTGATTTCTCGAATTTGGCAGACTGCGGATAGAAATAAAAGAAATTTGGTAAGCTTCTCT 2066
Db 1935 GTTATCACTAGAACTTTGGCAACAGCTGACAAAAACAGAAAGAAATTTGGCGCTTGAA 1994
QY 2067 GAAGATGGCAAGATAAAGATAATTTCCGATTAAGCGCTTACATCTCCAAATACATATC 2126
Db 1995 GAAGAAAAAGGCGATACGACAACTTCAAGATCAACCGCTTCTGTCTAAATACCAAT 2054
QY 2127 AACCCGCTTTGACCAACCGCGTGAAGAGTATATCGCTCTGTGGAAGGCGCAAGATC 2186
Db 2055 AACCCAGCGATCGCTCATGGATTAGCGAGTATGTAGGTTCAAGTAGAAGTGGGCAAGTG 2114
QY 2187 GCGACTTGGTGTGGAACTCTGCTTTTGGGCGTAAACCCCAAAATTCGTGATCAAA 2246
Db 2115 GCTGACTTGGTATTGTGGAGTCAGCAATTTTGGGCGTGAACCCCAACATGATCATCAA 2174
QY 2247 GCGGATGGTGGTCTTCTCTGAAATGGGCGATTTCTAACCGCTGTGTGCCCACTCCCCAA 2306
Db 2175 GCGGATTCATTGCGTTAAGCCAAATGGGCGATGCGAACTTCTATCTCCCTACCCACAA 2234
QY 2307 CCGGTTTATACCGCAATATTTTGGCATCAGCGCAAGCGGAAATTTGACACAGATC 2366
Db 2235 CCGGTTTATACAGAGAAATGTTTCGCTCATCATGGTAAAGCTAAATACGATGCAAAATC 2294
QY 2367 ACTTTTGGTTTCCAAAGTCGCTATGAAATGGGCGTGAAGAAAGCTGGGCTTTAGAGCGC 2426

Db 2295 ACTTTTGTGTCTCAAGCGCTTATGACAAAGGCAATTAAGAGAAATTAGGACTTTGAAAGA 2354
QY 2427 CAAGTTTCTACCGGTCAAAAACCTCCGTAAACATCACCAAGAAAGACTTCAAGTTCAACGAC 2486
Db 2355 CAAGTGTGTCGGTAAAAAATTCAGAGAAATATCACTAAAAAGACATGCAATTTCAACGAC 2414
QY 2487 AAAACGGCAAAATACACCGTCGATCCGAAAAACCTTCGAGGTCTTTGTAGATGGCAAACTC 2546
Db 2415 ACTACTGCTCACATTTGAAGTCAATCTCTGAAACTTACCATGTGTTTCGTGGATGGCAAGAA 2474
QY 2547 TGCACTCTTAAACCCACTCGCAAGTCCTCTAGCCAGCGCTACACTTTTCTTAGG 2604
Db 2475 GTAACTTTAAACAGCCATAAAGTAGAGCTTGCGCAACTCTTTAGCAATTTTCTTAGG 2532

RESULT 3

US-08-467-822-19
; Sequence 19, Application US/08467822
; Patent No. 5843460
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,822
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/447,177
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2619 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 31..36
; OTHER INFORMATION: /standard_name= "Shine-Dalgarno

| | | | | | | | | | |
|--|------|---|------|--|--|--|--|--|--|
| OTHER INFORMATION: sequence." | | | | | | | | | |
| FEATURE: | | | | | | | | | |
| NAME/KEY: misc feature | | | | | | | | | |
| LOCATION: 756..759 | | | | | | | | | |
| OTHER INFORMATION: /standard_name= "Shine-Dalgarno | | | | | | | | | |
| OTHER INFORMATION: sequence." | | | | | | | | | |
| US-08-467-822-19 | | | | | | | | | |
| Query Match 36.9%; Score 1062.6; DB 2; Length 2619; | | | | | | | | | |
| Best Local Similarity 65.5%; Pred. No.1e-263; | | | | | | | | | |
| Matches 1606; Conservative 0; Mismatches 804; Indels 41; Gaps 2; | | | | | | | | | |
| QY | 192 | ACAAAGGAGTAAATAGTGAAGTCAACCCAAAGAGCAAGAAAGTCTCTGTTATATATG | 251 | | | | | | |
| DB | 29 | ATAAGGAGTTAGGATGAAGTAAACGCTAAAGAACTAGACAAGTTAATGCTCCATTATG | 88 | | | | | | |
| QY | 252 | CGGGCGAAGTGGCTAGAAAGGGCAAGACGAGAGGCTTAAAGCTCAACCAACCCGAAGCCA | 311 | | | | | | |
| DB | 89 | CGGGCAGATTGGCGAAGAAACGCTTGGCGGTGGTGAACCTCAATTTACACCGAAGCGG | 148 | | | | | | |
| QY | 312 | TTGCTTACATTAGTGGCCCATATATATGACGAAGCGCGCTGGGAAAAAACCCTTGGCC | 371 | | | | | | |
| DB | 149 | TCGGCGCTATTAGCGGCGGTGATGGAAGAGCGCGTGATGGTAATAAAGCGTGGCGG | 208 | | | | | | |
| QY | 372 | AGCTTATGGAAGATGTCATGCACATTTTGTGAAAAAGATGAAGTAATGCCCCGGGTGGGTA | 431 | | | | | | |
| DB | 209 | ATTGTATGCAAGAGGCGAGGACTTGGCTTAAAAAAGAAATGTGATGGACGCGTAGCAA | 268 | | | | | | |
| QY | 432 | ATATGTTCCCGATCTAGTGTGTAGAACGCACTTTCTCTGTATGGTACGAACTTGTAACTG | 491 | | | | | | |
| DB | 269 | GCATGATTTCATGAAGTGGGATTTGAAGCTAACTTCCCGCATGGAACCAAGCTTGTAACTA | 328 | | | | | | |
| QY | 492 | TGAATGGCCCATCGAACACGATGAGCAGCTCAAGCGGGCGAGTGAATTTGGTTGCG | 551 | | | | | | |
| DB | 329 | TCACACTCCGGTAGAGGATAATGGCAAAATAGCCCCCGCGAGGTCTTCTTAA--AA | 385 | | | | | | |
| QY | 552 | ATAAAGACATCGAGCTCAATGCGAGCAAGAAAGTAAACCGAACTTGAGGTTACTAATGAAG | 611 | | | | | | |
| DB | 386 | ATAGGACATTACTATTAAACCGCGCAAGAGACCATTAGCTTGAAGTGAATAAAG | 445 | | | | | | |
| QY | 612 | GGCCTAAATCTTGCAATGCGGTAGCCATTTCCAATTTCTTTGAAAGCTAACAAAGCACATA | 671 | | | | | | |
| DB | 446 | CGCATCGTCTGTGCGAGTGGATCAATTTCCACTTCTTCCGAAGTGAATTAAGCTCTTGG | 505 | | | | | | |
| QY | 672 | AATTGATCTGTAAGAAAGCTATGGCAAGCCCTAGATATTCCTCTGCGCAACAGCTAC | 731 | | | | | | |
| DB | 506 | ACTTCGATCGCAAAAGCTTTTGCAAGCCCTAGACATTTGCATCTTGGAAACAGCGGTGC | 565 | | | | | | |
| QY | 732 | GCATTTGGGCGAGACAAACCGCAAGTGCAGTTGATTCTCTTGGTGGCAGTAAAAAG | 791 | | | | | | |
| DB | 566 | GCTTTGAACCCGGGAGGAAAAAGTGTGAACTCATTTGACATCGCGGGGAATTAAGCGCA | 625 | | | | | | |
| QY | 792 | TGATTGGCATGAACGGGCTTTGTGAATAACATCGCGGATGAACCGCCATAAAACATAAAGCGC | 851 | | | | | | |
| DB | 626 | TCTATGCTTTAATCTTCTTGTGGATCGCAAGCCGATGCCGATGGTAAAAAACTCGGCT | 685 | | | | | | |
| QY | 852 | TTGACAAAGGCGAAATCTCAGGATTT----- | 877 | | | | | | |
| DB | 686 | TAAAAACGCTTAAAGAAAAAGTTTGGGTGTGTAAACTTGGCTTGTGAAGCGCACTAAAG | 745 | | | | | | |
| QY | 878 | ----ATCAAGTAAGGAGACTCCCATGAATGAATAAGAAAAACAGAAATATGTAATACCTACG | 933 | | | | | | |
| DB | 746 | ATAAACAAATTAAGGAAAAACCATGAATAAGATTTTCAACGAAAGAAATATGTTCTATGTATG | 805 | | | | | | |
| QY | 934 | GACCCCAAGAGGCGATAAAGTGGCTTAGAGATACCGATCTTTGGGCGAAGCTAGAAC | 993 | | | | | | |
| DB | 806 | GTCCCACTACGGGGATCGTTTAGACTCGCGACACTGATTGATCTTAGAAGTGGAGC | 865 | | | | | | |
| QY | 994 | ATGACTATACCACTATGGCGAAGAACTTAAATTTGGCGGGGTAAAACTATCCGTGAGG | 1053 | | | | | | |
| DB | 866 | ATGATTGCACCACTTATGTTGAAGAGATCAAAATTTGGGGGGGTAAAAACTATCCGTGATG | 925 | | | | | | |
| QY | 1054 | GTATGGGTAGAGCAATAGCCCTGATGAAAAACACCCCTAGATTTAGTCATCACTAACCGCA | 1113 | | | | | | |

| | | | | | | | | | |
|----|------|---|------|--|--|--|--|--|--|
| DB | 926 | GGATGAGTCAAAACCAATAGCCCTAGCTCTTATGAATTAGATTGGTGTCTCACTAAACCCC | 985 | | | | | | |
| QY | 1114 | TGATTATCGACTACACCGGATTTTACAAAGCCGCACATTGGGATTTAAAAACGGCAAAATCC | 1173 | | | | | | |
| DB | 986 | TCATTGTGACTATACGGGCAATTTACAAAGCCGCACATTTGGGATTTAAAGACGGCAAGTTG | 1045 | | | | | | |
| QY | 1174 | ATGGCAATTGGCAAGGCGAGAAACAAGGACATGCAAGATGGCGTAAGCCCTCATATGTGTCG | 1233 | | | | | | |
| DB | 1046 | CAGGCATTGGCAAGGCGAGCAATAAGGACATGCAAGATGGCGTAGATAATAATCTTTTGGC | 1105 | | | | | | |
| QY | 1234 | TGGGTGTGGGACAGAACGACTAGCAGGGGAAGGTATGATTATTACGGCTGGGGGAATCG | 1293 | | | | | | |
| DB | 1106 | TAGGTCTCTCTACAGAGGCTTTGGCAGCTGAGGGCTTGATTTGAACCGCTGGTGGGCATCG | 1165 | | | | | | |
| QY | 1294 | ATTACACACACCCACTTTCTTCTCCACAACAATTCCTTACCCTCTAGCCCAATGGCCGTTA | 1353 | | | | | | |
| DB | 1166 | ATACGCATATTTCACATTATCTCTCCCCAACAAATCCCTACTGCTTTTTCAGCGGGGTTA | 1225 | | | | | | |
| QY | 1354 | CAACCATGTTTGGAGCGGCACAGGTCTGTAGATGGCAGCAATGCGACTTACTATCACTC | 1413 | | | | | | |
| DB | 1226 | CAACCATGATTGGAGGAGSCACAGGACCTGCGGATGGCAGCAATGCGACCACCATCACTC | 1285 | | | | | | |
| QY | 1414 | CGGCAAAATGGAACTTTGCACCCGATGTTGCGCGCAGCAGAAAGTATTTCTATGAATGTGG | 1473 | | | | | | |
| DB | 1286 | CGGACGCGCTAAATCTTAAAAAGTATGTTGCGTGCAGCCGGAAGAAATACGCCATGAATCTAG | 1345 | | | | | | |
| QY | 1474 | GCTTTTTCGGCAAAAGCAATAGCTCTAGCAAAAAACAATCTGTAGAACCAAGTAGAAGCGG | 1533 | | | | | | |
| DB | 1346 | GCTTTTTCGGCTAAGGGGAATGTCTTACGAACCCCTTTTACGCGATCAGATTGAAGCAG | 1405 | | | | | | |
| QY | 1534 | GCGCATTTGGTTTAAATTTGCAATGAAGACTGGGGCAACAACCAACAGTGGGATCGATCACT | 1593 | | | | | | |
| DB | 1406 | GGGCGATTTGGTTTAAATTCACGAAGACTGGGGAAGCACACTGCGAGCTATTTCACCACT | 1465 | | | | | | |
| QY | 1594 | GCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATACAGTCAATG | 1653 | | | | | | |
| DB | 1466 | GCCTCAATGTGCGCCGATGAATACGATGTGCAAGTGGCTATCCACACCGATACCTTAAACG | 1525 | | | | | | |
| QY | 1654 | AGGAGGTTATGTAGATGACACCTTAAATGCAATGAACGGGCGCGCATCCATGCCTTACC | 1713 | | | | | | |
| DB | 1526 | AGGCGGCTGTGTAGAAGACACCTTAGAGGCGATTTGCGCGGCGCACCATCCATACCTTCC | 1585 | | | | | | |
| QY | 1714 | ACATTTGAGGAGCGGTGGAGGACACTCACCTGATGTTATCACATGGCAGGCGAGTCA | 1773 | | | | | | |
| DB | 1586 | ACATGGAAGGGCTGGGGTGGACAGCTCCAGATGTTATCAAATTTGCGAGGGAATTTA | 1645 | | | | | | |
| QY | 1774 | ATATTCTACCTCTCCACACCCCACTATTTCCTCTATACCATTAATACGGTTGCGAAGC | 1833 | | | | | | |
| DB | 1646 | ACATTTCTACCCGCTCTACTAACCCGACCATTTCTTTTACCAAAAAACACTGAAGCCGAGC | 1705 | | | | | | |
| QY | 1834 | ACTTAGACATGTCTATGACATGCCACCTAGACAAACGATCCCGGAGGATTTTACAAT | 1893 | | | | | | |
| DB | 1706 | ACATGGACATGTTAATGTGTGCCACCACTTGGATAAAAAAGTATCAAGGAAGATGTGCAGT | 1765 | | | | | | |
| QY | 1894 | TTTCTCAAAGCGTATCCGCCCGGCTCTATCCGCTGAAGATGTCTCCATGATATGG | 1953 | | | | | | |
| DB | 1766 | TTGCGGATTCGAGGATTCGCCCCCAAACTATCGGGCTGAAGACCAACTCCATGACATGG | 1825 | | | | | | |
| QY | 1954 | GTGTGATCGCATGACAAAGCTCGGATTCGAAAGCAATGGGGGTGCGAGGCGAAGTGAATTC | 2013 | | | | | | |
| DB | 1826 | GGATCTTTCTATCACAGCTCCGACTCTCAGGCTATGGGACGCGTAGGCGAGGTGATCA | 1885 | | | | | | |
| QY | 2014 | CTCGAACTTGGCAGACTCGGATAAGAAATAAAAAAGAAATTTGGTAAGCTTCTCGAAGATG | 2073 | | | | | | |
| DB | 1886 | CAGGCATTGGCAGACAGCAGACAAAAACAAAAAGAGTTTGGGCGCTTGAAGAGAGGAAA | 1945 | | | | | | |
| QY | 2074 | GCAAGATAACGATTAATTTCCGCATTAGCGCTACATCTCCAAATACACTATCAACCCCG | 2133 | | | | | | |
| DB | 1946 | AAGGCGATAACGACAACTTCCGCATCAAAACGCTACATCTCTAAATACACCATCAACCCCG | 2005 | | | | | | |
| QY | 2134 | CTTTGACCCACGCGGTGAGGAGTATATCGGCTCTGTGGAAGAGGGCAAGATCGCCGACT | 2193 | | | | | | |

Db 2066 GGATCGCGCATGGGATTTCTGACTATGTGGGCTCTGTGGAAGTGGGCAATACGCCGACC 2065
QY 2194 TGGTGTGTGGGAATCTCGCTCTTTTGGCGTAAACCCCAAAATCGTGATCAAAAGCGGTA 2253
Db 2066 TCGTGCTTTGGAGTCGGCTTCTTTGGCATTAAGCCCATATGATTTAAAGGGCGAT 2125
QY 2254 TGGTGTCTTCTCTGAAATGGGCGATTCTAAACCGCTCTGTGCCACTTCCCAACCGGTTT 2313
Db 2126 TTATTGCGCTCTCTCAATGGGCGATGCAATGTCGTCTATTCCCAACCCCTCAGCCGCT 2185
QY 2314 ATTACCGCAAAATGTTTGGGCATCGGCAGCAAGCGGAAATTTGACACGAGATCACTTTG 2373
Db 2186 ATTACCGTGAATGTTTGGACACCATGGGAAACAAATTCGACACCAATATCACTTCG 2245
QY 2374 TTTCCAAAGTCGCTATGAAATGGGCTGAAAGAAAGCTGGGCTTAGAGCGCAAGTTC 2433
Db 2246 TGTCCCAAGCGCTTACAGGCGAGGATCNAAGAGAACTAGGGCTAGATCGCGCGCAC 2305
QY 2434 TACCGGTCAAAATCTCGGTAAACATCACCAAGAAAGACTTCAAGTTCAACGCAAAACGG 2493
Db 2306 CGCCAGTGAAGAACTGTCGAATATCACTAAAGAGGACCTCAAAATCAACGATGTGACCG 2365
QY 2494 CAAAATACCCGTCGATCCGAAACCTTCGAGTCTTTGTAGATGGCAAACTCTGCACCT 2553
Db 2366 CACATATTGATGTCAACCCCTGAAACCTATAAGGTGAAAGTGGATGGCAAGAGGTAACT 2425
QY 2554 CTAACCCACCTCGCAAGTGCCTCTAGCCCGAGCTACACTTTCTTCTAGG 2604
Db 2426 CTAACGACGAGTGAATTTAGCCTAGCGCACTTTATATTTGTTCTAGG 2476

RESULT 4

US-08-432-697-19
; Sequence 19, Application US/08432697
; Patent No. 6248330
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TREATMENT OF INFECTIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; COMPOSITIONS
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patencin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-00000
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2619 base pairs
; TYPE: nucleic acid

; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 31..36
; OTHER INFORMATION: /standard_name= "Shine-Dalgarno
; OTHER INFORMATION: sequence."
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 755..759
; OTHER INFORMATION: /standard_name= "Shine-Dalgarno
; OTHER INFORMATION: sequence."
; US-08-432-697-19
Query Match 36.9%; Score 1062.6; DB 3; Length 2619;
Best Local Similarity 65.5%; Pred. No. 1e-263;
Matches 1606; Conservative 0; Mismatches 804; Indels 41; Gaps 2;
QY 192 ACAAGGAGTAATAGTGGAACTCACCCCAAGAGCAAGAAAGTTCTTTTATATTATG 251
Db 29 ATAAGGAGTTTAGGATGAACCTAACGCTTAAGAACTAGACAAGTTAATGCTCCATTATG 88
QY 252 CGGCGAAGTGGCTAGAAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCCGAAGCCA 311
Db 89 CGGCGAGATTGGCAGAAAGACGCTTGGCGCTGTGTGAACTCAATTACACCGAAGCGG 148
QY 312 TTGCTTACATTTAGTCCCATATTATGACAGAGCGCGCTGGAAAAAAGCTTGGCCCC 371
Db 149 TCGCGCTCATTTAGCGGCGTGTGATGAAAAGGCGCTGATGTAATAAAGCGTGGCGG 208
QY 372 AGCTTATGGAAGAGTGCATGCACCTTTTGAAGAAAGATGAAGTATGCCCGGGTGGTA 431
Db 209 ATTGATGCAAGAGGCGAGACTTGGCTTAAAGAAAGAAATGATGACGCGTAGCAA 268
QY 432 ATATGTTCCGATCTAGGTGTAGAACCACTTTCTGTAGGTAGCAAACTTTGTAACGTG 491
Db 269 GCATGATTCATGAAGTGGGATTTGAAGTAACTTCCCGATGGAACCAAGCTTTGTAACTA 328
QY 492 TGAATTTGCCCATCGAACCCAGATGACACTTCAAAGCGGCGGAAGTAAATTTGGTGG 551
Db 329 TCCACACTCCGGTAGAGATATGGCAATTTAGCCCCGGGAGGTCTTCTTAAA--AA 385
QY 552 ATAAAGACATCGAGTCAATGCAGCAAGAAAGTAAACCGAACTTGAGGTACTAATGAAG 611
Db 386 ATGAGGACATTACTATTAAACCGCGCAAGAAAGCCATTAGCTTGAAGTGAATAAAG 445
QY 612 GGCCTAAATCCTTGCATGTGGGTAGCCATTTCCACTTTCTTTGAAGCTAACCAAGGCACTAA 671
Db 446 GCGATCGTCTGTGCGAGTGGGATCACATTTCCACTTCTTCGAAGTGAATAAGCTCTTGG 505
QY 672 AATTGATCGTGAAGAAAGCCATATGCAAAAGCTAGATATTTCCCTCTGCAACACGCTAC 731
Db 506 ACTTCGATCGCGCAAAAGCTTTTGAACGCGCTAGACATTTGCACTCTGGAACGCGGTGC 565
QY 732 GCATTTGGGAGGACAAACCCCGAAAGTGCAGTTGATTTCTCTGTGTGCGAGTAAAGAAAG 791
Db 566 GCTTTGAACCCGGGAGGAAAAAGTGTGGAACCTCATTTGACATCGCGGGAATAGCGCA 625
QY 792 TGATTGGCATGAACGGGCTTGTGAATAACATCGCGGATGAACGCCATAAACAATAAGCGCG 851
Db 626 TCTATGGCTTAAATTTCTTTGGTGGATCGCAAGCGGATGCCGATGTTGTAATAAAGCTCGGCT 685
QY 852 TTGACAGGCGAAATCTCAGGATTT----- 877
Db 686 TAAACCGCGCTAAAGAAAAAGGTTTGGGTCTGTAAGCTGCGGTGTGTAAGGACATAAG 745
QY 878 ----ATCAAGTAAGGAGACTCCCATGAAATGAAGAAACAAAGAAATATGTAATACCTACG 933
Db 746 ATAAACAATAAGGAAAAACCATGAAAAAGATTTTCGAAAAAGATATGTTTCTATGTAAG 805
QY 934 GACCCACCAAGGCGGATAAAGTGGCTTTAGGAGATACCGATCTTTGGGCGAGAGTAGAAC 993

Db 806 GTCCCACTACCGGGATCGTGTAGACTCGCGGACACTGATTTGATCTTAGAAGTGAGC 865
Qy 994 ATGACTATACACCTATGGCGAAGAACTTAAATTTGGCGGGGTAAAACTATCCGTGAGG 1053
Db 866 ATGATTGCACCACTTATGGTGAAGAGATCAAATTTGGGGCGGTAAAACTATCCGTGATG 925
Qy 1054 GTATGGGTAGAGCAATAGCCCTGTATGAAAAACCCCTAGATTTAGTCACTCAACCGGA 1113
Db 926 GGATGAGTCAAAACCAATAGCCCTAGCTCTTATGAATTAGAATTTGGTGTCTCACTAACGCC 985
Qy 1114 TGATTATCGACTACACCGGATTTACAAGCGACATTTGGATTTAAAGCGGCAAAATCC 1173
Db 986 TCATTGTGGACTATACGGGCAITTTACAAGCCGACATTTGGATTTAAAGAGCGGCAAGATTG 1045
Qy 1174 ATGGCATTTGCAAGGAGGAAACAAGGACATGCAAGATGCGTAAGCCCTCATATGCTG 1233
Db 1046 CAGGCATTGCAAGCGAGGCAATAGGACATGCAAGATGGCGTAGATAAATCTTTGCG 1105
Qy 1234 TGGGTGTGGCACAGAACTAGCAGGGAAGGTATGATTTATPACCGCTGGGGGAATCG 1293
Db 1106 TAGGTCTGCTACAGAGGCTTTGGCAGCTGAGGGCTTGAATTTAAACCGCTGGTGGCATG 1165
Qy 1294 ATTACACACCCAGTCTCTTCTCCACAAATTCCTCTACCGCTCTAGCCCAATGGCGTTA 1353
Db 1166 ATACGCATATTCATCTTTATCTCTCCCAACAAATCCCTACTGCTTTTGCAGCGGGGTTA 1225
Qy 1354 CAACCATTTTGGAGCGGCACAGGTCCTGTAGATGCGCAAGATGCGCACTACTATCACTC 1413
Db 1226 CAACCATGATTGGAGGAGGCACAGGACCTCGGATGGCAAGATGGCAACCACTCACTC 1285
Qy 1414 CGCGCAAAATGGAACTTGCACCGCATTTGCGCGCAGCAGAGAGTATTTCTATGAATGTGG 1473
Db 1286 CCGGACGGCTATCTAAAGATATGTTGCGCGAGCGGAAGATACGCCATGATCTAG 1345
Qy 1474 GCTTTTGGCAAGGCAATAGCTCTAGCAAAAAACAACTTGTAGAACAACTAGTAAGCGG 1533
Db 1346 GCTTTTGGCTAAGGGGAATGTGTCTTACGAAACCTCTTTACGCGATCAGATTGAAGCAG 1405
Qy 1534 GCGGATTTGTTTAAATTCATGAGACTGGGGCAACACCAAGTGGCATCACT 1593
Db 1406 GGGCGATTTGTTTAAATCCAGAGACTGGGGAAGCAACCTTGCAGCTATTCAACACT 1465
Qy 1594 GCTTGAGCGTGGCAGATGAATACGATGCAAGTTTGTATCCACACCGATACAGTCAATG 1653
Db 1466 GCCTCAATGTCGCGATGAATACGATGCGAGTGGCTATCCACACCGATACCTTAACG 1525
Qy 1654 AGCAGGTTATGATGACACACCTTAAATGCAATGAAACGCGCGCCCATCCATCCCTACC 1713
Db 1526 AGCGGGCTGTGTAAGAGACACCTTAGAGCGGATTTGCGGGCGCACCATCATACCTTCC 1585
Qy 1714 ACATTGAGGAGCGGTGGAGGACACTCACTGATGTTATCACCATGGCGAGCGAGCTCA 1773
Db 1586 ACATGAAGGGGTGGGGGTGGACACGCTCCAGATGTTATCAAAATGGCAGGGGAATTTA 1645
Qy 1774 ATATTCTACCTCTCCACACCCCACTATTCCTATACCATTAATACGTTTGCAGAAC 1833
Db 1646 ACATTCTACCGGCTCTACTAAACCGGACCAATCTCTTCACAAAAACACTGAAGCCGAGC 1705
Qy 1834 ACTTAGACATGCTCATGACATGCCACCACTAGACAAACGATCCGCGAGGATTTACAAT 1893
Db 1706 ACATGACATGTTAATGGTGTGCCACCACTTGTGATAAAGTATCAAGGAGATGTGCAAT 1765
Qy 1894 TTTCTCAAAGCCGTATCCGCCCGGCTTATCGCGGCTGAAGATGTGCTCATGATATGG 1953
Db 1766 TTGCCGATTGAGGATTCGCCCGCCCAAACTATCGCGGCTGAAGACCAACTCCATGACATGG 1825
Qy 1954 GTGTGATCGGATGACAGCTCGATTGCGAAGCAATGGGGCTGCGAGCGGAAGTATTC 2013
Db 1826 GGATCTTTTCTATCACAGCTCCGACTCTCAGGCTATGGACGCGTAGGCGAGGTATCA 1885
Qy 2014 CTCGAACTTGGCAGACTGCGGATAAGAAATAAAAAAGAAATTTGGTAAGCTTCTCAAGATG 2073
Db 1886 CAGGCATTTGGCAGACGACAGACAAAAAACAANAAGAGTTTGGCGCTTGAAGAGGAA 1945

Qy 2074 GCAAAGATAACGATAATTTCCGCATTAAGCGCTACATCTCCAAATACACTATCAACCCCG 2133
Db 1946 AAGCGATAACGACAACTTCGCGATCAAAACGCTACATCTCTAAATACACCATCAACCCCG 2005
Qy 2134 CTTTGACCCACGGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGGCAGATCGCCGACT 2193
Db 2006 GGATCGCGCATGGGATTTCTGACTATATGTGGGCTCTGTGGAAGTGGCAAAATACGCCGACC 2065
Qy 2194 TGGTGGTGTGAATCCTGCCTTTTGGCGTAAAAACCCAAATCGTGATCAAAAGGCGGTA 2253
Db 2066 TCGTGTCTTTGAGTCCGCTTTCTTTGGCATTAAGCCCAATATGATTATTAAGGGCGAT 2125
Qy 2254 TGGTGGTCTTCTGAAATGGCGGATTTAAACGCGTCTGTGCCCCACTCCCCAACCGGTTT 2313
Db 2126 TTATTGGCTCTCTCAAAATGGCGATGCCAATGCGTCTATTCCACCCCTCAGCCGCTCT 2185
Qy 2314 ATTACCGGAAATGTTTGGGATCAAGCGAGCGGCAAAATTTGACACCGATCACTTTTG 2373
Db 2186 ATTACCGTGAATGTTTGGACACCATGGGAAAAACAAATTCGACCAATATCACTTTTCG 2245
Qy 2374 TTTCCAAGTCGCTTATGAAATGCGTGAAGAAAGCTGGCTTAGAGCGCAAGTTC 2433
Db 2246 TGTCCCAAGCGCTTACAAGGAGGATCAAGAAAGAACTAGGGCTAGATCGCGGCGAC 2305
Qy 2434 TACCGCTCAAAAACCTGCGCTAACATCACCAAGAAAGACTTCAAGTTTCAACGACAAACCG 2493
Db 2306 CGCCAGTCAAAAACCTGTCGAATATCACTAAAAGGACTCAAAATTCACGATGTGACCG 2365
Qy 2494 CAAAAATCACCGTGCATCCGAAAACCTTCGAGGTCTTTGTAGATGGCAAACTCTGCACCT 2553
Db 2366 CACATTTGATGTCAACCTGAAACCTATAAGGTGAAAGTGGATGGCAAGAGGTAACTT 2425
Qy 2554 CTAAACCCACTCGCAAGTGCCTCTAGCCGAGCGCTACACTTTCTCTAGG 2604
Db 2426 CTAAAGCAGCAGATGCAATTTAGCCCTAGCGCACTTTATAATTTGTCTAGG 2476

RESULT 5

US-08-466-248-19
; Sequence 19, Application US/08466248
; Patent No. 6258359
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,248
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/447,177
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2619 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 31..36
; OTHER INFORMATION: /standard name= "Shine-Dalgarno"
; OTHER INFORMATION: sequence."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 756..759
; OTHER INFORMATION: /standard name= "Shine-Dalgarno"
; OTHER INFORMATION: sequence."
US-08-466-248-19

Query Match 36.9%; Score 1062.6; DB 3; Length 2619;
Best Local Similarity 65.5%; Pred. No. 1e-263;
Matches 1606; Conservative 0; Mismatches 804; Indels 41; Gaps 2;

QY 192 ACAAGAGTAATAGGTGAACCTCACACCCCAAGAGCAAGAAAGTCTTGTGTATATATG 251
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 252 CGGGCGAAGTGGCTAGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCCGGAAGCCA 311
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 89 CGGCAGATTGGCAGNAGAACGCTTGGCGGTGTGAACTCAATTACCCGAAGCGG 148
QY 312 TTGCTTACATAGTCCCATATATATGACGAGCGCGGTGGAAGAAACCCGTTGCC 371
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 149 TCGCGCTCATAGCGCGGTGATGGAAGAGCGCGTGTATGATTAAGAGCGTGGCG 208
QY 372 ACCTTATGAAGAGTGCATGCACTTTTGAAGAAAGATGAAGTAATGCCCGGGGGTGA 431
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 209 ATTTGATGCAAGAGCGGACCTTGGCTTAAAGAAAGAAATGTGATGAGCGGTAGCAA 268
QY 432 ATATGGTCCCATCTAGGTGTAGAGCCACCTTTCTGTATGCTACGAACTTGTAACTG 491
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 269 GCATGATTCATGAAGTGGGATTTGAAGCTTAACCTTCCCGATGGAACCAAGCTTGTAACTA 328
QY 492 TGAATTTGGCCCATCGAACCCAGATGACACTTCAAAGCGGCGCAAGTGAATTTGGTTGCG 551
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 329 TCCACACTCGGTAGAGGATATGGCAATATAGCCCGCGGAGGCTTCTTAAA ---AA 395
QY 552 ATAAGACATCGAGCTCAATGACGAGCAAGAGTAACCGAACTTGAGGTTACTAATGAAG 611
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 386 ATGAGGACATTAATTAACCGCGGCAAGAGCCATTAGCTTTGAAGTGAATTAAG 445
QY 612 GGCCTAAATCCTTGCATGGGTAGCCATTTCCACTTCTTTGAAGCTAACAGGCACTAA 671
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 446 GCGATCGCTGTGCAAGTGGGATCACTTTCCACTTCTTGAAGTGAATTAAGCTCTTGG 505
QY 672 AATTTCGATCGTGAAGAGCTATGGCAACCGCTAGATATTCCTCTTGGCAACACGCTAC 731
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 506 ACTTCGATCGCGCAAAAGCTTTTGAACACCGCTTAGACATTCGATCTGGAACAGCGGTG 565
QY 732 GCATTGGGCGAGGACAAACCCGCAAGTGCAGTTGATTCCTCTTGGTGGCAGTAAAAAAG 791
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 566 GCTTTGAACCCCGGGAGGAAAAAGTGTGAACTCAATTGATCGCGGGGAATAAGCGCA 625
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

792 TGATTGGCATGAACGGGCTTGTGAATAACATCCCGGATGAACCGCATATAAATAAAGCGC 851
626 TCTATGGCTTTAATTTCTTGGTGGATCGCCAGCGATGCCGATGGTAAATAAACTCGGCT 685
852 TTGACAAGGGGAATCTCAGGATTT----- 877
686 TAAACCGCGCTAAAGAAAAAGGTTTTTGGGCTCTGTAACCTCGGTTGTGAAGCGACTAAAG 745
878 ----ATCAAGTAAGGAGACTCCCATGAAATGAAAAAACAAGAAATATGTAATACCTACG 933
746 ATAAACAATTAAGAAAAACCATGAAAAAGATTTCAAGAAAAAGATATGTTCTATGTAAG 805
934 GACCCACCAAGGCGATATAAGTCGCTTAGAGATACCGATCTTTTGGGCGAAGTAGAAC 993
806 GTCCCACTACCGGGGATCGTTTAGACTCGCGCACACTGATTGATCTTAGAAGTGGAGC 865
994 ATGACTATACCACTATGGCGAAGAACTTAAATTTGGCGGGGTAAAACTATCCGTGAGG 1053
866 ATGATTGCACCACTTATGGTGAAGAGATCAAAATTTGGGGCGGTAAAACTATCCGTGATG 925
1054 GTATGGTCTAGAGCAATAGCCCTGATGAAAAACACCCCTAGATTATTAGTCATCACTAACGCGA 1113
926 GGATGAGTCAACCAATAGCCCTAGCTCTTATGAATTAGATTGTGCTCACTAACGCC 985
1114 TGATTATCGACTACACCGGATTTTCAAAAGCCGACATTTGGGATTTAAAAACCGCAAAATCC 1173
986 TCATTGTGGACTATACGGGCATTTTCAAAAGCCGACATTTGGGATTTAAAGACGCGAAGTTG 1045
1174 ATGGCATTTGGCAAGGCGGAAACAAGGACATGCAAGATGGCGTAAAGCCCTCATATGTTGTCG 1233
1046 CAGGCATTTGGCAAGGCGGCAATAAGGACATGCAAGATGGCGTAGATAAATAATCTTTGCG 1105
1234 TGGGTGTGGCAGAGAGCACTAGCAGGGAAGTATGATTATACCCTGGGGGAATCG 1293
1106 TAGGTCTCTACAGAGGCTTTGGCAGCTGAGGCTTGAATTGAACCGCTGGTGCCATCG 1165
1294 ATTCAACACACCCACTCTTCTTCCCAACAATTTCCCTACCGCTCTAGCCCAATGGCGTTA 1353
1166 ATACGATATTCACTTTATCTCTCCCAACAATACTCTACTGCTTTTGGCAGCGGGTTA 1225
1354 CAACCATTTTGGAGGCGGACAGGCTCTGTAGATGGCAAGAAATGCGCACTACTATCACTC 1413
1226 CAACCATGATTGGAGGAGGACAGGACCTCGCGATGGCAAGAAATGCGCACTACTCACTC 1285
1414 CGGGCAATGGAATTCACCGCATGTTGCGCGCAGCAGAGAGTATTTCTATGATGTGG 1473
1286 CCGGACGCGCTAATCTAAAAAAGTATGTTGCGTGCAGCCGAAGAAATACGCCATGAATCTAG 1345
1474 GCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAAACAATTTGTAGACAAGTAGAAGCGG 1533
1346 GCTTTTGGCTAAGGGAATGTGTCTTACGAAACCTCTTTACGCGATCAGATTGAAGCAG 1405
1534 GCGCGATTTGTTTTTAAATTTGCATGAAGACTGGGGCAACAACCAAGTGCAGTCACTACT 1593
1406 GGGGATTTGTTTTTAAATCCACGAAGACTGGGGAAGCACACCTGCGAGCTATTCACTACT 1465
1594 GCTTGAGCGTGGCAGATGAATACGATGTGAAGTTTGTATCCACACCGATACAGTCAATG 1653
1466 GCCTCAATGTGCGCGATGAATACGATGTGAAGTGGCTATCCACACCGATACCTTAAAG 1525
1654 AGGCGGTTATGTAGATGACACCTTAATGCAATGAACGGGCGGCCATCCATGCCTACC 1713
1526 AGCGGGCTGTGTAGAAAGACACCTTAGAGGCGATTGCCGGCGGCACCATCCATACCTTCC 1585
1714 ACATTGAGGAGCGGGTGGAGGACACTCACCTGATGTTATCACCATGCGAGCGAGCTCA 1773
1586 ACATTGAGGGCTGGGGTGGACACGCTCAGATGTTATCAAAATGGCAGGGGAATTTA 1645
1774 ATATTCTACCTCTCTCAACCCCGCCCTATTCCTTATACATTAATTAATCGGTTCAGAAC 1833
1646 ACATTCTACCGGCTCTACTTAACCCGACCATTCCTTTCAACCAAAAACTGAAAGCGGAGC 1705
```

```
QY 1834 ACTTAGACATGCTCATGACATGCCACCACTAGACAAACGATCCGCGAGGATTTACAAT 1893
Db 1706 ACATGACATGTTAAATGTTGGTGGCCACCCTTGGATAAAAGTATCAAGGAAGATGTGCAGT 1765
QY 1894 TTCTCNAACCGTATCCGCGCGGCTCTATCGGGCTGAAGATGTGCTCATGATATGG 1953
Db 1766 TTCCGATTCAGAGATTTCGCGCCCAAACTATCGCGGTGAAGACCAAACTCCATGACATGG 1825
QY 1954 GTGTGATCGGATGACAAGCTTCGATTCGAAGCAATGGGGCGTGCAGGGCAAGTGATTC 2013
Db 1826 GGATCTTTCTATCACAGCTCCGACTCTCAGGCTATGGACGGTGAAGGAGTGATCA 1885
QY 2014 CTCGAACCTTGGACAGCTCGGATAAGATAAAGAAATTTGGTAAGCTTCTCTGAAGATG 2073
Db 1886 CACGCACTTCGACAGACAGACAACAAAACAAAAGAGTTTGGCGCTTGAAGAGGAAA 1945
QY 2074 GCAAGATTAAGTAATTTCCGATAAAGCGCTACATCTCCAATACACTATCAACCCGG 2133
Db 1946 AAGCGGATAACGACAACTTCGCGATCAACGCTACATCTCTAAATACACCATCAACCCCG 2005
QY 2134 CTTTGACCCAGCGGTGAGCGAGTATATCGGCTCTGTGGAGAGGGCAAGATCGCGACT 2193
Db 2006 GGATCGCGCATGGATTTCTGACTATGTGGCTCTGTGGAGTGGGCAAAATACGCCGACC 2065
QY 2194 TGGTGTGTGGAATCTCTGCTTTTGGCGTAAACCCCAAAATCGTGATCAAAAGCGGTA 2253
Db 2066 TCGTGCTTTGAGTCCGGCTTTCTTTGGCATTAAAGCCATATGATTTAAGGGCGGAT 2125
QY 2254 TGGTGTGTTCTGAAATGGGCGATTCTAAACGGCTCTGTGCCACTCCCAACCGGTTT 2313
Db 2126 TTATTGGCTCTCTCAATGGGCGATGCCAATGCGTCTATTCCACCCCTCAGCCCGTCT 2185
QY 2314 ATTACCGGAAATGTTTGGGCATCAGCGGAGGGAATTTGACACGAGCATCACTTTTG 2373
Db 2186 ATTACCGTGAATGTTTGGACACCATGGGAAAACAAAATTCGACACCAATATCACTTTCG 2245
QY 2374 TTTCCAAAGTCGCTATGAAATGGCGTGAAAGAAAAGCTGGGCTTAGAGCGCCAAAGTTC 2433
Db 2246 TGTCCAAGCGGCTTACAGGCGAGGATCAAGAGAACTAGGCTAGATCGCGCGAC 2305
QY 2434 TACCGGTCAAAAATCGCGTAAACATCAACCAAGAAAGACTTTCAAGTTCAACGACAAAACGG 2493
Db 2306 CGCCAGTGA AAAACTGTGCAATATCACTAAAAGGACCTCAAATTCACGATGTGACCG 2365
QY 2494 CAAAAATCACCGTCGATCCGAAACCTTCGAGGTCTTTGTAGATGGCAAACTCTGCACT 2553
Db 2366 CACATATTGATGTCAACCCCTGAAACCTTATAAGGTGAAAGTGGATGGCAAGAGGTAACCT 2425
QY 2554 CTAAACCCACTCGCAAGTGCCTCTAGCCGAGCGCTACACTTTCTCTAGG 2604
Db 2426 CTAAAGCAGCAGATGAATTGAGCTAGCGCAACTTTATAATTTGTTCTAGG 2476
```

RESULT 6

```
US-09-431-705-1
; Sequence 1, Application US/09431705
; Patent No. 6585975
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Londono-Arcila, Patricia
; APPLICANT: Freeman, Donna
; TITLE OF INVENTION: Use of salmonella vectors for
; TITLE OF INVENTION: vaccination against helicobacter infection
; FILE REFERENCE: 06132/060001
; CURRENT APPLICATION NUMBER: US/09/431,705
; CURRENT FILING DATE: 1999-11-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 4824
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
```

```
; OTHER INFORMATION: includes sequences from Helicobacter pylori,
; OTHER INFORMATION: Salmonella typhimurium, and Escherichia coli
; NAME/KEY: CDS
; LOCATION: (2)...(31)
; NAME/KEY: CDS
; LOCATION: (41)...(61)
; NAME/KEY: CDS
; LOCATION: (65)...(799)
; NAME/KEY: CDS
; LOCATION: (803)...(2512)
; NAME/KEY: CDS
; LOCATION: (2516)...(2692)
; NAME/KEY: CDS
; LOCATION: (2696)...(2896)
; NAME/KEY: CDS
; LOCATION: (2900)...(3322)
; NAME/KEY: CDS
; LOCATION: (3326)...(3385)
; NAME/KEY: CDS
; LOCATION: (3389)...(3406)
; NAME/KEY: CDS
; LOCATION: (3410)...(3466)
; NAME/KEY: CDS
; LOCATION: (3470)...(3598)
; NAME/KEY: CDS
; LOCATION: (3602)...(3661)
; NAME/KEY: CDS
; LOCATION: (3665)...(3697)
; NAME/KEY: CDS
; LOCATION: (3701)...(3769)
; NAME/KEY: CDS
; LOCATION: (3773)...(3817)
; NAME/KEY: CDS
; LOCATION: (3821)...(3844)
; NAME/KEY: CDS
; LOCATION: (3848)...(3889)
; US-09-431-705-1
```

Query Match 36.7%; Score 1058.6; DB 3; Length 4824;

Best Local Similarity 65.7%; Pred. No. 1.4e-262;

Matches 1606; Conservative 0; Mismatches 799; Indels 38; Gaps 3;

```
QY 197 GAGTAATAGGTGAAACTCACACCCAAAGAGCAAGAAAGTTCTTGTATATTATCGGGC 256
Db 77 GAGATCTCCATGAATACTCACCCCAAGAGTTAGATAAGTTGATGCTCCACTACGCTGA 136
QY 257 GAAGTGGCTAGAAAGCGCAAGAGAGGGCTTAAAGCTCAACCAACCGAAGCCATTGCT 316
Db 137 GAATTTGGCTAAAAACGCAAGAAAGGCACTTAAGCTTAACTATGTAGAAGCAGTAGCT 196
QY 317 TACATTAGTCCCATATATTATGACGAGCGCGCTGAAAAAAGCCGTTGCCAGCTT 376
Db 197 TTGATTAGTCCCATATATTATGGAAGAGCAGAGCTGTGTAAGAAAGACTGCGGCTGAAATG 256
QY 377 ATGGAAGAGTGCATGCACCTTTTGAAGAAAGATGAAGTAATGCCCGGGTGGGTAATATG 436
Db 257 ATGCAAGAGAGGCGCATCTCTTTTAAACCATGATGATGATGGATGGCGTGGCAAGCATG 316
QY 437 GTTCCCGATCTAGGTGTAGAAAGCCACCTTTCTCTGATGGTACGAAACTTTGTAACTGTGAAT 496
Db 317 ATCCATGAAGTGGGTATTGAAGCGATGTTTCTGATGGGACTTAAACTCGTAACCGTGCAAT 376
QY 497 TGGCCCATCGAAACGAGATGAGCATTCAAGGGGCGAAGTGAATTTGGTTGGCATAAA 556
Db 377 ACCCCTATTGAGGCAATGGTAAATTAGTCTCTGCTGAGTTG---TTCCTAAAAAATGAA 433
QY 557 GACATCGAGCTCAATGAGGCAAGAGTAACCGAACTTGGAGTTTACTTAATGAAGGCCT 616
Db 434 GACATCACTATCAAGAAAGCAAAAGCCGTTAGCGGTGAAAGTTTAAATAATTTGCGCAC 493
QY 617 AAATCCTTTGATGTGGGTAGGCATTTCCACTTTCTTTGAAGCTTAAACAGGCACATAAATTC 676
Db 494 AGACCGGTTCAATCGGCTCACACTTCCATTTCTTTGAAGTAAGATAGATGCCTAGACTTT 553
```

QY 677 GATCGTGAAGAACCTTATGGCAACGCCCTAGATATTTCCCTCTGGCAACACGCTACGCAAT 736
D 678 |||||
D 679 |||||
QY 680 GACAGAGAAAACCTTTTCGGTAAACGCTTAGACATTTGCGAGCGGACAGCGTAAAGATTT 613
D 681 |||||
QY 682 GGGGAGGACAAACCCGCAAGTGCAGTGTATTCCTCTTGGTGGCAGTAAAGAAAGTAT 796
D 683 |||||
D 684 GAGCCTGGCGAAGAAAATCCGTAGAAATTTGATGACATTTTGGCGGTAAACAGAAATCTTT 673
QY 685 GGCATGAACGGCTTGTGATTAACATCGCGGATGAACGCCATAAACATAAAGCGCTTGAC 856
D 686 |||||
D 687 GATTTAACGCATTTGGTGTATGACAGACAAACGCAAGCAAAATTTGCTTTTACAC 733
QY 688 AAGGCGAAATCTCACGATTTT-----ATCAAGT 884
D 689 |||||
D 690 AGAGCTAAAGAGCGTGTGTTTCATGCGCTTAAAGCGATGACAACTATGTAAACAAAT 793
QY 691 AAGGAGACTCCCATGAAA--ATGAAAAACAAGATATGTAAATACCTACGAGACCCACC 941
D 692 |||||
D 693 AAGGAGTAAAGAAATGAAAAAGATTTAGCAGAAAAGATATGTTCTATGTATGCTCTACT 853
QY 694 AAGGCGATAAAGTGGCTTGGCGATACAGACTTGATCGCTGAAGTAGAACATGACTAC 913
D 695 |||||
QY 696 AAGGCGATAAAGTGGCTTGGCGATACAGACTTGATCGCTGAAGTAGAACATGACTAC 1001
D 697 |||||
D 698 AAGGCGATAAAGTGGCTTGGCGATACAGACTTGATCGCTGAAGTAGAACATGACTAC 913
QY 699 ACCACCTATGGCGAAGAACTTAAATTTGGCGGGTAAACCTATCCGTGAGGGTATGGGT 1061
D 700 |||||
D 701 ACCATTTATGGCGAAGAGCTTAAATTTGGCGGGTAAACCTATCCGTGAGGGTATGGGT 973
QY 702 CAGAGCAATAGCCCTGATGAAACACCCCTAGATTTAGTCACTATCAACGCGATGATTC 1121
D 703 |||||
D 704 CAATCTAACCAACCTAGCAAGAGAGTGGATTTAAATTTATCACTAACCGCTTTAACTCGTG 1033
QY 705 GACTACACCGGATTTTACAAAGCGGACATTTGGATTTAAACCGGCAAAATCCATGGCAT 1181
D 706 |||||
D 707 GATTACACCGGATTTTAAAGCGGATTTGGATTTAAAGATGCGCAAAATCCGTGGCAT 1093
QY 708 GGCAGGAGGAAACAAAGACATGCAAGATGGCGTAAAGCCCTCATATGGTCTGGTGGTG 1241
D 709 |||||
D 710 GGTAAAGCGGTAAACAAAGACATGCAAGATGGCGTTAAACCAATCTTAGCGTAGGTCT 1153
QY 711 GGCACAGAGCACTAGCAGGGGAGGTATGATTTATACCGCTGGGGAAATCGATTCACAC 1301
D 712 |||||
D 713 GCTACTGAAGCCTTAGCCGGTGAAGTTTGATCGTAAAGCGCTGGTGGTATTTGACACAC 1213
QY 714 ACCACTTCTTCTCCACAACTTCCCTACCGCTCTAGCCAAATGGCGTTACACCATG 1361
D 715 |||||
D 716 ATCCACTTCAATTTACCCCAACAAATCCCTACAGCTTTTTCAGAGCGGTGTAACAAACCAT 1273
QY 717 TTTGGAGGCGCACAGGTCTGTAGATGGCAAGATGGCAATGCGACTACTATCACTCCGGGCAA 1421
D 718 |||||
D 719 ATTGGTGGTGAACCGGTCTGTGATGGCACTTAATGCGACTACTATCACTCCAGGCGA 1333
QY 720 TGGAACTTGCACGCACTTTTGGCGGAGCAGAGAGATTTCTATGAATGTGGGCTTTTG 1481
D 721 |||||
D 722 AGAAATTTAAATGGATGCTCAGAGCGGTGAAGATATTTCTATGAATTTAGTTTCTTG 1393
QY 723 GGCAGAGCAATAGCTCTAGCAAAAACAACTTTGTAGAAACAGTGAAGCGGGCGCAT 1541
D 724 |||||
D 725 GCTAAAGGTAAACGCTTCTAACGATGCGAGCTTAGCCGATCAAAATTTGAAGCGGTGCGAT 1453
QY 726 GGTTTTAAATTTGATGAAGACTGGGCGACAAACCAAGTGGGATCGATCACTGCTTGAGC 1601
D 727 |||||
D 728 GGTTTTGAATTTCAAGAACTGGGCGACACCTCTCTCTGCAATCAATCATGCGGTAGAT 1513
QY 729 GTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATACAGTCAATGAGGAGGT 1661
D 730 |||||
D 731 GTTGGCGACAAATACGATGTGCAAGTGGCTATCGCCACACACACTTTGAATGAAGCGGT 1573
QY 732 TATGTATGACACCTTAATGAATGAACGGCGGCGCATTCATGCTTACCAATTTGAG 1721
D 733 |||||
D 734 TGTGTAGAGACACTATGCTGCTATTGCTGGAGCGCACTATGACACACTTTTCCACACTGAA 1633

QY 1722 GGAGCGGTGGAGGACACTCACCTGATGTTATACCATGGAGCGAGCTCAATATTCTTA 1781
D 1723 |||||
D 1724 GCGCTGGCGCGGACACAGCTCTCTGATTTATTTAAAGTAGCCGTGAACAACAATTTCT 1693
QY 1725 CCCTCTCCACCAACCCCACTATTTCCCTATACCAATTAACGTTTGCAGACACTTAGAC 1841
D 1726 |||||
D 1727 CCGCTTCCACTAACCCCACTCCCTTTCCCGTGAATACAGAAAGCAGACATGAC 1753
QY 1728 ATGCTCATGACATGCCACCACTAGACAAAACGCGATTCGCGAGGATTTTCAAAATTTCTCAA 1901
D 1729 |||||
D 1730 ATGCTTATGTTGGTGGCACACTTGGATTAAGCAATTAAGAAAGATGTTCACTCGCTGAT 1813
QY 1731 AGCGTATTCGCGCGGCTCTATCGCGCTGAAAGATGTCTTCAATGATATGGTGTGATC 1961
D 1732 |||||
D 1733 TCAAGGATTCGCGCTCAACCAATTTGCGCTGAAGACACTTTTGCATGACATGGGATTTTC 1873
QY 1734 GCATGACAGCTCGGATTCGCAAGCAATGGGGCGTGCAGCGGAGTATTCCTCGAAT 2021
D 1735 |||||
D 1736 TCAATCAACAGTTTCTGACTCTCAAGCGATGGCGCTGTGGGTGAAGTTATCACTAGAAT 1933
QY 1737 TGGCAGACTCGGATTAAGATTAAGAAAGATTTTGGTAAAGCTTCTCTGAAGATGGCAAGAT 2081
D 1738 |||||
D 1739 TGGCAACAGCTGACAAAAACAGAAAGATTTTGGCGCTTGAAGAAAGAAAAAGCGAT 1993
QY 1740 AACGATATTTCCGCAATTAAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACC 2141
D 1741 |||||
D 1742 AACGCAACTTTCAGATCAAGCGCTACTTGTCTAAATACACCAATTAACCCAGCGATCGCT 2053
QY 1743 CAGCGGTGAGCGATATATCGGCTCTGTGGAAGAGGCAAGATCGCGACTTGTGGTG 2201
D 1744 |||||
D 1745 CATGGGATTAAGCGAGTATGAGTTTCAAGTGAAGTGGCAAGTGGCTGACTTGGTATTTG 2113
QY 1746 TGGAACTCTCTCTTTTGGGTAAACCCCAAAATCGTGATCAAGGCGGTATGTTGGTGC 2261
D 1747 |||||
D 1748 TGGAGTCCAGATTTCTTTGGCGTAAACCCCAAGATGATCAAGGCGGATTCATGCG 2173
QY 1749 TTCTCTGAAATGGCGGATTTTAAACGCTCTGTGCGCACTCCCCCAACCGGTTTATACCGC 2321
D 1750 |||||
D 1751 TTAAGCCAAATGGCGATGCGAAGCTTCTATCTCTACCCCAACCGGTTTATACAGA 2233
QY 1752 GAAATGTTGGGCAATCGGCAAGCGGAAATTTGACACCGATCACTTTTGTGTTTCCAAA 2381
D 1753 |||||
D 1754 GAAATGTTGCTCATCATGTGTAAAGCTAAATACGATGCAATGCAAAACATCACTTTTGTGCTCAA 2293
QY 1755 GTGCGCTATGAAATGGCGTGAAGAAAGCTGGCTTAGAGCGCAAGTTCTACCGGTG 2441
D 1756 |||||
D 1757 GCGGCTTATGACAAAGCAATTAAGAAAGATTAAGACTTTGAAGACAAAGTTGTCGGTGA 2353
QY 1758 AAAAACTGGCGTAAACATCAACCAAGAAAGCTTCAAGTTCAACGACAAACCGCAAAATC 2501
D 1759 |||||
D 1760 AAAAAATTGCAAAATATCACTAAAGAAAGACATGCAATTCACGACACTACCGCTCACAT 2413
QY 1761 ACCGTTCGATCGAAAACTTGGAGGTCTTTGTAGATGGCAAACTCTGCACCTCTAAACCC 2561
D 1762 |||||
D 1763 GAAGTCAATCTCGAACTTACCATGTGTTGCGTGAAGTGAAGAAAGTAACTTCTAAACCA 2473
QY 1764 ACCTCGCAAGTGCCTTAGCCCGCGCTACACTTTCTTCTAGG 2604
D 1765 |||||
D 1766 GCGAATAAAGTGAAGTGGCGCAACTCTTTTAGCATTTTCTAGG 2516

RESULT 7

US-09-431-705-19
; Sequence 19, Application US/09431705
; Patent No. 6585975
; GENERAL INFORMATION:
; APPLICANT: Kleantous, Harold
; APPLICANT: Londono-Arcila, Patricia
; APPLICANT: Freeman, Donna
; TITLE OF INVENTION: Use of salmonella vectors for
; TITLE OF INVENTION: vaccination against helicobacter infection
; FILE REFERENCE: 06132/060001
; CURRENT APPLICATION NUMBER: US/09/431,705

| | | | |
|----|------|--|------|
| Qy | 781 | CAGTAAAAAAGTGTGATTCGTCATGAACGGGCTTTGTGAATAACATCGCGATGAACGCCATAA | 840 |
| Db | 914 | TGAACGCAAGTGTATCGTTAAATAATGTAAACGA----- | 948 |
| Qy | 841 | ACATAAGCGCTTGAACAAGCGGAATCTCACGGATTTATCAAGTAAGGAGACTCCCAATGA | 900 |
| Db | 949 | --ATGGATCAGTTTCAAAATGGGAAAAGAAAATGAGTT----- | 985 |
| Qy | 901 | AAATGAAAAACAAGATATGTAAATACCTACCGACCCACCAAGGCGGATAAAGTCGCCT | 960 |
| Db | 987 | CGATGTCTCGAAAAGCAATATACGGGATATGTTTGGACCAACTGTGCGGCGACGCCATTCGTT | 1046 |
| Qy | 961 | TAGGAGATACCGATCTTTTGGSCAGAAGTAGAACAATGACTATACCACTATATGGCGGAAGAAC | 1020 |
| Db | 1047 | TGGCAGATTCAGAAATGTTTATCGAAATTTGAAAAGGACTATACACGTATGGAGATGAGG | 1106 |
| Qy | 1021 | TTAAATTTGGCGGGGTAAAACTATCCGTGAGGGTATGGGTTCAGAGCAAT---AGGCCTG | 1077 |
| Db | 1107 | TAAAGTTTGGCGGCGCAAGGTGATCCGAGATGGAATGGGCGAGCATCTTTGGCGCAAA | 1166 |
| Qy | 1078 | ATGAAACACCCCTAGATTTTATGTATCATCTAACCGGATGATTTATCGACTACACCGGAGTTT | 1137 |
| Db | 1167 | GGCATGAATCGTCGATCTCGTATTAACAAATGCGATTATTTGTGTGATTAACACAGGTATTT | 1226 |
| Qy | 1138 | ACAAAGCGACATTTGGGATTTAAAAACCGCAAAATCCATGCGCATTTGGCAAGCGCAGGAACA | 1197 |
| Db | 1227 | ATAAAGCAGATATCGGCATATAAAGATGGAATGATTTGCTCCATAGGAAAAGCGGGGAACC | 1286 |
| Qy | 1198 | AGGACATGCAAGATGGCGTAAAGCCCTCATATGGTGTGGGTGTGGGCACAGAAGCACTAG | 1257 |
| Db | 1287 | CGTTGTTAATGAGCGGGGTG-----ATATGGTGATTTGAGCAGCAAGAGTCTATAG | 1340 |
| Qy | 1258 | CAGGGGAAGGTATGATTAATACCGCTGGGGGAATTCGATTACACACCCACTTCCTTTCTC | 1317 |
| Db | 1341 | CCGCAGAAGGATGATTTGTGACAGCGGAGGAATAGATGCTCATATTTCACTTTATTTGCC | 1400 |
| Qy | 1318 | CACAAATTTCCCTACCGCTCTAGCCAAATGGCGTTTAAACATGTTTGGAGGCGGCACAG | 1377 |
| Db | 1401 | CTCAGCAATTCGAAACCGCTCTTGATCGGGGTGTGACCACTATGATTTGGCGGAGGAACAG | 1460 |
| Qy | 1378 | GTCTGTAGATGGCACGAATTCGCACTACTATCACTCCGGGCAATGGAATTGCAACCGCA | 1437 |
| Db | 1461 | GACCCGCTCAGGCACAAATGCCACTACTTTGTACACCGGGGCCCTTGGAAATATCCATCGTA | 1520 |
| Qy | 1438 | TGTTTGGCGCAGCAGAAGAGTATTTCTATGAATGTGGCTTTTTTGGGCAAAAGGCAATAGCT | 1497 |
| Db | 1521 | TGCTTTAAGCAGCGCAAGAAATCCCGATAAACTTTGGCTTTTTTAGGAAAGGGAACCTGTT | 1580 |
| Qy | 1498 | CTAGCAAAAAACAACCTTGTAGAACAGTGTAGAGCGGGCGGATTTGGTTTTTAAATTTGCATG | 1557 |
| Db | 1581 | CAGATGAGGCTCCTTTTAAAGGAACAAATTTGAAGCGGAGCGGTGGGATTTAAGCTTCAAG | 1640 |
| Qy | 1558 | AAGATCGGGCAACAAACCAAGTCCGATTCGATCATCTGCTTGAGCGTGGCAGATGAATACG | 1617 |
| Db | 1641 | AAGATTTGGGATCGACGCGCGCGCTATTGTATACATCTTTTGAAGTGGCGGATCGATATG | 1700 |
| Qy | 1618 | ATGTGCAAGTTTTGTATCCACACCGATACAGTCAATGAGGCAGGTATATGTAGATGACACCC | 1677 |
| Db | 1701 | ATGTGCAAGTAGCGATTTATACAGACACTTTTAAATGAAGCGGATTTTGTGAGGATACTTT | 1760 |
| Qy | 1678 | TAAATGCAATGAACGGCGCGCCATCATGCTTACCACATTTGAGGGAGCGGTGGAGGAC | 1737 |
| Db | 1761 | TGAAGCCATAGCGGTGAGTGATTTATATCTATCATACAGAAGGGGCTGGCGGGGAC | 1820 |
| Qy | 1738 | ACTCACCTGATGTTATCACCATGGCAGCGAGCTCAATATTTCTACCTCTCCACCAACC | 1797 |
| Db | 1821 | ATGCTCCGATATATAAAGCGCGCGCTTCCCGAATATTTTGCCTTCTTCCACGAATC | 1880 |
| Qy | 1798 | CCACTATTCCTCATACCATTAATACGGTTGAGAAACCTTAGACATGCTCATGACATGCC | 1857 |
| Db | 1881 | CAACTCCAGCTTATACTATCAATACTTTTGAAGAGCAATTTAGATATGTTAATGTTTGGC | 1940 |

RESULT 9
US-08-967-513-1
; Sequence 1, Application US/08967513
; Patent No. 5783436
; GENERAL INFORMATION:
; APPLICANT: Robert P. Hausinger
; TITLE OF INVENTION: Mutant Urease and Method
; TITLE OF INVENTION: of Use For Determination
; TITLE OF INVENTION: of Urea
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ian C. McLeod
; STREET: 2190 Commons Parkway
; CITY: Okemos
; STATE: Michigan
; COUNTRY: USA

```
/
/
/  ZIP: 48864
/  COMPUTER READABLE FORM:
/  MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb
/  MEDIUM TYPE: storage
/  COMPUTER: IBM Compatible
/  OPERATING SYSTEM: MS-DOS
/  SOFTWARE: Wordperfect 5.1
/  CURRENT APPLICATION DATA:
/  APPLICATION NUMBER: US/08/967,513
/  FILING DATE: 11-NOV-1997
/  CLASSIFICATION: 435
/  PRIOR APPLICATION DATA:
/  APPLICATION NUMBER: 08/687,645
/  FILING DATE: July 26, 1996
/  ATTORNEY/AGENT INFORMATION:
/  NAME: Ian C. McLeod
/  REGISTRATION NUMBER: 20,931
/  REFERENCE/DOCKET NUMBER: MSU 4.1-309
/  TELECOMMUNICATION INFORMATION:
/  TELEPHONE: (517) 347-4100
/  TELEFAX: (517) 347-4103
/  TELEX: No. 5783436e
/  INFORMATION FOR SEQ ID NO: 1:
/  SEQUENCE CHARACTERISTICS:
/  LENGTH: 2400
/  TYPE: nucleotides
/  STRANDEDNESS: Single
/  TOPOLOGY: Linear
/  MOLSCULE TYPE:
/  DESCRIPTION: cDNA
/  HYPOTHETICAL: NO
/  ANTI-SENSE: NO
/  ORIGINAL SOURCE:
/  ORGANISM: Klebsiella aerogenes
/  STRAIN: CG253
/  INDIVIDUAL ISOLATE:
/  CELL TYPE: N/A
/  FEATURE:
/  NAME/KEY: cDNA encoding mutant urease
/  NAME/KEY: 'H219Q
/  LOCATION: Modification at position 1312 to
/  LOCATION: glutamine
/  IDENTIFICATION METHOD: Sequencing
/  OTHER INFORMATION:
/
/  US-08-967-513-1
/
Query Match 24.1%; Score 693.4; DB 2; Length 2400;
Best Local Similarity 57.1%; Pred. No. 1.7e-168;
Matches 1385; Conservative 0; Mismatches 971; Indels 69; Gaps 4;

QY 195 AGGAGTAATAGGTGAACCTCACACCCAAAGAGCAAGAAAGTTCTTGTATATATGCGG 254
Db 20 AAGAGAACGTATGGAACCTGACCCCGAGAAAAAGACAAGCTGTGTGCTTTACCGCG 79
QY 255 GCGAAGTGGCTAGAAAGCGCAAGAGAGGCGCTTAAAGCTCAACCAACCGCAAGCCATTG 314
Db 80 CGCTGTGGCGGAGCGTGGCTGGCCGCGGCGCTGAAGCTCAACTATCCGAGTCCGTGG 139
QY 315 CTTACATTAGTGCCTATTTATGAGACAGCGCGCGTGGAAAAAACCCTTTGCCGAGC 374
Db 140 CCTGATCAGCGCTTTATTATGGAAGGCGCTCGGAGCGG---CAAAGCGTGGCTCGC 196
QY 375 TTATGGAAGATGCATGCACCTTTTGAAGAAAGATGAAGTAATGCCCGGGGTGGTAATA 434
Db 197 TGATGGAGAAAGGCGCTCACTGCTTACCGCGAGCAGGTGATGAGGAGGCGTCCCGGAAA 256
QY 435 TGGTTCCTCCGATCTAGGTGTAGAGCCACCTTTCTCGATGTCGAAACTTGTAACTGTGA 494
Db 257 TGATCCCGGATATCCAGGTGGAAGCACCCTTTCCCGGAGCGCTCGAAGCTGGTCAACGTT 316
QY 495 ATTGGCCCATCGAACCAGATGAGCACTTCAAAGCGGCGCAAGTGAATTTGGTTGGGATA 554
Db 317 ACAACCCGATTATCTGAGGTAGCCCATGATCCCGGTGAATATCACGTTAAGCCCGGTC 376
```

```
QY 555 AAGACATCGAGCTCAATGCGAGGCAAGAAAGTAACCGAACTTGAGGTTACTAATGAAGGC 614
Db 377 ---AGATAGCCCTGAATACCGCGCGGCAACCTGTGCGGTGGTCTTTGAGAACACACGCG 433
QY 615 CTAAATCCTTGGATGTGGGTAGCGATTTCCACTTCTTTTGAAGCTAAACAGGCACATAAAT 674
Db 434 ATCGGCGGATTTCAGGTGCGGTTCGCACTACCATTTTCGCCGAGGTTAACCGCGCTGAAGT 493
QY 675 TCGATCGTGAAGAAAGCCTATGGCAACCGCTAGATATTTCCCTCTGGCAACACGCTTACGCA 734
Db 494 TCGACCGTTCAGAGCGCGCGGCTATCGCTGAAATATCCCGCGCGGACACGCGGTACGCT 553
QY 735 TTGGGCGAGGACAAACCCGCAAAAGTGCAGTTGATTCCTCTTGTGGCAGTAAAAAAGTGA 794
Db 554 TTGAACCCGCGCAGAAACGCGAGGTGAGCTGCTGCGCGGTTCGCCGCGTCAACCGCGCT 613
QY 795 TTGGCATGAACCGGCTTTGAAATAACATCGCGGATGAACCGCATAAACATAAAGGCTTG 854
Db 614 TCGGCTTTCGCGCGGAGGTTCATGGGCGCTCTCGAGGTAAACGAT----- 657
QY 855 ACAAGGCGAATCTCAGGATTTATCAAGTAAGGAGACTCCCATGAAATGAAAAACAA 914
Db 658 -----GAGTAATATTTACGCCAG 676
QY 915 GAATATGTAATACCTACGGACCCACCAAGGCGATAAAGTGCCTTTAGGAGATACCGAT 974
Db 677 GCCTATGCCGATATGTTTCGGCCCCACCGTCGGCGACAAGGTGCGCTTGGCAGATACCGAG 736
QY 975 CTTTGGGCGAAGTAGAACATGACTATACCACTTATGGCGAAGAACTTAAATTTGGCGCG 1034
Db 737 CTGTGATCGAGGTGGAGGACGATTTGACCACTTACGGGGAAGAGGTCAAAATTCGGCGCG 796
QY 1035 GGTAAACTATCCGTAGGCTATGGGTACAGAGCAATAGCCTGTATGAAACACCTAGAT 1094
Db 797 GGCANAAGTATCCGCGACGCGATGGCGCAGGACAGATGCTGGCGCGCGACTGTGTGCGAC 856
QY 1095 TTAGTCATCACTAACGCGATGATTATCGACTACACCGGGATTTACAAAGCCGACATTTGGG 1154
Db 857 CTGCTGCTCAACAACGCTGTATCGTCGTCGATCGTGGGGATCGTTAAGCCGATATCGCG 916
QY 1155 ATTAAAAACGGCAAAATCCATGCGATTGGCAAGCAGGAAACAAAGGACATCAAGATGCG 1214
Db 917 GTGAAGGAGCGCGGATCTTTCGCGCATCGCAAAAGCGGCAACCCGACATCCAGCCCAAC 976
QY 1215 GTAAGCCCTCATATGCTGCTGGGTGGGCGACAGACGACTAGCAGGGAAGGTATGATT 1274
Db 977 GTCAAC-----ATCCCCCATCGCGCGCTCGCAGCGAAGTGATCGCGCGCAAGAAAAAT 1030
QY 1275 ATTACCGCTCGGGGAATCGATTTCACACCCACCTTCTCTTCTCCACAACAATTCCTTACC 1334
Db 1031 GTCACCCCGCGGGATTCGATACCCATATTCCTCTGGATCTGTCTCGCAGCAGCGGGAAGAG 1090
QY 1335 GCTCTAGCCCAATGGCTTACAAACCATGTTTGGAGCGCGCACAGGTCTCTGTAGTGGCAGC 1394
Db 1091 GCGTGTCTCTGGGCTGACCAACCATGTCGCGCGGCGCACCCGCGCGCGCGCGGCGCACC 1150
QY 1395 AATGCGACTATCATCTCCGGGCAATGGAACCTTGACACGCGATGTTGGCGCGCAGAGAA 1454
Db 1151 CATGCCACCACTGCACCCCGCGCGCTGTGTATATCTCACGCATGCTCGAGCGCGCGCAC 1210
QY 1455 GAGTATTTCTATGATGTGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAATTT 1514
Db 1211 AGCCTGCGGTCAATATCGGCGCTGTGGGCAAGGAAACGTTTCTACGCGGATGCCCTG 1270
QY 1515 GTAGAACAAAGTAGAAGCGGCGCATTTGGTTTAAATTTGAATGAAGCTGGGGCACAACA 1574
Db 1271 CGCGAGCAGGTGGCGGAGCGGTTTATTTGGCTGAAGATCCAAGAGGACTGGGGCGCCACC 1330
QY 1575 CCAAGTCGATCGATCATCTGCTGAGGTGGCAGATGAATACGATGTGCAAGTTTGTATC 1634
Db 1331 CCGCGCGCGATCGACTGTGCGTTAAACCGTCGCCGATGAATGGACATCCAGGTGCGCCCTG 1390
```


| | | | | | | | |
|----|------|---|------|----|------|---|------|
| QY | 375 | TTATCGAAGAGTGCATGCACTTTTGAAGAAAGATGAAGTAATGCCCCGGGTGGTAAATA | 434 | QY | 1455 | GAGTATTATGAATGTGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAATTT | 1514 |
| Db | 197 | TGATGGAGGAGGCGCTCAGCTCTGACCGCGAGCAGGTGATGGAGGCGCTCCCGGAAA | 256 | Db | 1211 | AGCTCCCGTCAATATCGGCTCTGCGCAAGGAAACGTTTCTCAGCCGATGCCCCTG | 1270 |
| QY | 435 | TGGTTCCTGATAGTGTGAGAGCACCCTTCTGATGATGATGAGAACTTGAACGTGA | 494 | QY | 1515 | GTAGAACAGTAGAAGCGGCGGATTTTAAATTGCATGAAGACTGGGGGACACACA | 1574 |
| Db | 257 | TGATCCGGGATATCCAGGTGCAAGCCACCTTCCCGACGGCTCGAAGCTGGTCACCGTTC | 316 | Db | 1271 | CGGAGAGGTGGCGGAGCGCTTATTTGGCTGAAGATCAAGAGGACTGGGGCGCCACC | 1330 |
| QY | 495 | ATTGGCCCATGAAACAGATGAGCACTTCAAAGCGGCGAAGTGAATTTGGTTGGATA | 554 | QY | 1575 | CCAAGTGCATCGATCACTGCTTGAAGCTGGCAGATGAATACGATGTGCAAGTTGTATC | 1634 |
| Db | 317 | ACAAACCGATTATCGAGGTAGCGCATGATCCCCGGTGAATATCACGTTAAGCCCGTTC | 376 | Db | 1331 | CGGCGCGGATCGACTGTGGTTAAACGTCGCGATGAATGGACATCCAGGTGCGCCCTG | 1390 |
| QY | 555 | AAGACATCGAGCTCAATGCGAGCAAGAAAGTAAACCGAATCTGAGGTACTTAATGAAGGCG | 614 | QY | 1635 | CACACGATACAGTCAATAGGCGAGGTATGTAGATGACACCTTAAATGCAATGAACGGG | 1694 |
| Db | 377 | ---AGATAGCCCTGAATACCGGCGGCGCAACCTGTGCGGTGGTCTGTGAGAACCAACCGCG | 433 | Db | 1391 | CACAGGACACCTTGAATGAATCCGGTTTGTGTGAAGACACCTCGCGCCATCGGCGGG | 1450 |
| QY | 615 | CTAAATCTTGTGATGTTGGTAGGCTATTCGACTTCTTTTGAAGCTAACAAGGCACTAAAT | 674 | QY | 1695 | CGGCGCATCCATGCTTACCACTTGAAGGAGCGGTGGAGGACACTCACCTGATGTTATC | 1754 |
| Db | 434 | ATCGGCCGATTGAGTTCGCTTACCTAGCACTTTCGCCGAGGTAAACCGGCGCTGAAGT | 493 | Db | 1451 | CGACCATCCACACTTCCATACCGAAGGGCGCGCGGCCCATGCGCCGGACATCATC | 1510 |
| QY | 675 | TCGATCTGAAAGGCTATGCGCAAGCCTAGATATTCCTCTGTCGCAACACGCTACGCA | 734 | QY | 1755 | ACCATGCGAGCGAGCTCAATATTCTACCTCTCTCCACACCCGCCACTATTTCCTTATACC | 1814 |
| Db | 494 | TCGACCGTCAGAGCGCGGCTATCGCTCTGAATATCCCGGCGGACGCGCGGTACGCT | 553 | Db | 1511 | ACGCGTGGCGCCACCGAATATTTGCGCTGCTGTCACCAACCCAAACGCTGCCCTACACC | 1570 |
| QY | 735 | TTGGGCGAGCAAAACCCGAAAGTGCAGTTGATTCCTCTTGTGTGCAGTAAAAAAGTGA | 794 | QY | 1815 | ATTAAATACGTTGCGAGAACCTTAGACATGCTCATGACATGCCACCACTAGACAAACGC | 1874 |
| Db | 554 | TTGAACCCGCGCAGAAACGGAGTGCAGTGTGTGCTTCCGCGCTACCGCGCGCTCT | 613 | Db | 1571 | CTCAACACCATCGATGAACATCTCGATATGCTGATGCTGCGCACCAACCTCTGACCCGGAC | 1630 |
| QY | 795 | TTGGCATGAACGGGCTTTGTAATAACATCGCGATGAACGCCATAAACATAAAGCGCTTG | 854 | QY | 1875 | ATCGCGAGGATTACAAATTTCTCAAAGCGGTATCGCCCCGGCTCTATCGCGCTGAA | 1934 |
| Db | 614 | TCGGCTTCGCGCGAGGTTCATGGGCCCTCTGGAGGTAAACGAT----- | 657 | Db | 1631 | ATCGCGAGGACTGGGCTTTTGGCGAGTTCGCGCATTCGCGGGAACCATCGCTGCGGAA | 1690 |
| QY | 855 | ACAAGCGAAATCTCAGCGATTATCAAGTAAGGAGTCCCATGAAATGAAGAAACAA | 914 | QY | 1935 | GATGTCTCCTCATGATGCGGTGATCGCGATGACAAAGTCCGATTCGCAAGCAATGGG | 1994 |
| Db | 658 | -----GAGTAAATATTTACGCCAG | 676 | Db | 1691 | GAGTGTGACGATCTCGGCGCTTCTCGCTACCTCTCCGATTCGAGGCCATGCGC | 1750 |
| QY | 915 | GAATATGTAAATACCTACGAGCCCAACAAAGCGATATAAGTGGCTTAGAGATACCGAT | 974 | QY | 1995 | CGTGACGGGAGTGTCTCGAACTTGGCAGATCGCGATAGAATAAAAAAGAAATTT | 2054 |
| Db | 677 | GCTATGCGGATATGTTCCGCCCAACGCTCGGCGACAGGTGCGCTGCGCAGATACGAG | 736 | Db | 1751 | CGCGTGGGGAAGTATTCTCCGACCTGGCAGGTGGCGCATCGCATGAAGTGCAGCGC | 1810 |
| QY | 975 | CTTTGGCGAGAGTAGAAATGATGATACACCTATGGCGAAGAACTTAAATTTGGCGCG | 1034 | QY | 2055 | GGTAAGCTTCTGAAAGTGGCAAGATATAAGTAATTTTCGCATTAAGCGCTACATCTCC | 2114 |
| Db | 737 | CTGTGATCGAGTGGAGGACGATTTGACCACTACGCGGAGAGGTCAAATTTGGCGCGC | 796 | Db | 1811 | GGAGCGTGGCGGAGGAGACCGGGGATAACGACAACTTCCGCGTGAAGCGCTACATCGCC | 1870 |
| QY | 1035 | GGTAAATCTCCGTGAGGATAGGTCAGAGCAATAGCCCTGATGAAGAACCCCTAGAT | 1094 | QY | 2115 | AAATACATATCAACCCGCTTTGACCCACGCGCTGAGCGATATATCGGCTCTGTGAA | 2174 |
| Db | 797 | GGCAAAAGTATCCGCGACGCGATGGGCGAGGACAGATGCTGGCCCGGACTGTGTGAC | 856 | Db | 1871 | AAATACACCATCAACCGGCGTGACCCAGCGCATCGCACAGAGTCGGATCCATTTGAG | 1930 |
| QY | 1095 | TTAGTCACTAAACCGGATGATTATCGACTACACCGGATTTTACAAAGCGGACATTTGG | 1154 | QY | 2175 | GAGGCAAGATCGCCGACTTGGTGTGGAATCTCTGCTTTTGGCGTAAAAACCCAAA | 2234 |
| Db | 857 | CTGGTCTCACCACCGGTTGATCTGATCTCTGGGGATCGTTAAGGCCGATATCGGC | 916 | Db | 1931 | GTGGTAAAGCTGGCTGACTCTGCTGCTGGTTCACAGGCTTCTTCGGCGTGAACCGGCC | 1990 |
| QY | 1155 | ATTAAACCGGAAATTCATGTCATTTGGCAAGCGAGGAAACAGGACATGCAAGATGGC | 1214 | QY | 2235 | ATCGTATCAAAAGCGGATGTTGTTCTCTGAAATTTGGCGGATTTTAAACGCTCTGTG | 2294 |
| Db | 917 | GTGAAGGACGCGCGATCTTCGCCATCGCAAGCGCGCAACCCCGACATCCAGCCCAAC | 976 | Db | 1991 | ACCGTATCAAAAGCGGATGATCGCCATCGCGCGATGGCGGATATCAATGCTCTATT | 2050 |
| QY | 1215 | GTAAGCCCTCATATGTTGTTGGTGTGGGACAGAGACATAGCAGGGAAGGTATGATT | 1274 | QY | 2295 | CCCACTCCCCAAACGGGTTTATACCGGCAATTTTGGGCATCAGGCAAGCGGAAATTT | 2354 |
| Db | 977 | GTCAAC-----ATCCCCATCGCGCTGCGAGCGAAGTATCGCCCGCGAAGGAAAAAT | 1030 | Db | 2051 | CCGACCCGCGCGGTCACCTACCGCCGATTTTGGCGGCTGGGCGCGCGCCCAT | 2110 |
| QY | 1275 | ATTACCGCTGGGGAAATCGATTACACACCCACTTCTCTTCTCCACAACTTTCCCTACC | 1334 | QY | 2355 | GACACGAGCATCTTTTGTTCAAAGTCCCTATGAAAAATGGCGTGAAGAAAAAGCTG | 2414 |
| Db | 1031 | GTACCCCGCGCGGATCGATACCCATATTCATGATCTGTCCGAGCAGCGCGGAAGAG | 1090 | Db | 2111 | CATGCGCTCACCCTTCTGTGCGAGCGCGGCGGCGCATGCGGTTGCGGAGCGGCTG | 2170 |
| QY | 1335 | GCTCTAGCAATGGCGTTACAACTATTTGGAGGGGCGCACAGGTCTCTGATGAGCAGC | 1394 | QY | 2415 | GGCTTAGAGCGCCAAAGTTCTACCGGTCAAAAATTCGCGTAAACATCACCAAGAAAGCTTC | 2474 |
| Db | 1091 | GGCTGTCTCTGGCGTGACCACTATGTTGGCGGCGCACCGCGCGCGCGCGGACCC | 1150 | Db | 2171 | AACCTGCGCAGCGGATCGCGCTGTGAAGGCTGCGGTACGGTGCAGAAAGCCGACATG | 2230 |
| QY | 1395 | AATGCGACTATATCATCTCGGGGAAATGAACTTGACCGCATTTGTCGCGCAGCAGAA | 1454 | QY | 2475 | AAGTTCAACAGCAAAAACGGCAAAATCACCGTGCATCGGAAACCTTCGAGGTCTTTGTA | 2534 |
| Db | 1151 | CATGCCAACCACTGCACCCCGGCGCGTGTATATCTCAGGCATGCTGCGAGGCGCGAC | 1210 | Db | 2231 | GTGCACAAACAGTCTGCGAGCCCTAAACATCCGTCGACGCCAGACCTATGAGGTGCGGGT | 2290 |
| | | | | QY | 2535 | GATGGCAAACTCTGCACTCTTAACCCACCTCGCAAGTGGCTCTAGCCCCAGCGCTACACT | 2594 |

| | | | |
|---|--------|--|--------|
| Db | 2291 | GATCGCAACTTATACACAGCGAGCGGACAGAGTTCGCGGATGCGCAACGATATTTT | 2350 |
| Qy | 2595 | TTCTTCTAGGCACAATGCCCTTT | 2619 |
| Db | 2351 | CTGTTTAAAGGAGCGGATGCTTT | 2375 |
| RESULT 11 | | | |
| US-09-557-884-1/c | | | |
| : Sequence 1, Application US/09557884 | | | |
| : Patent No. 6506581 | | | |
| : GENERAL INFORMATION: | | | |
| : APPLICANT: Fleischmann et al. | | | |
| : TITLE OF INVENTION: The Nucleotide sequence of | | | |
| : the Haemophilus influenzae Rd Genome, Fragments | | | |
| : Thereof, and Uses Thereof | | | |
| : NUMBER OF SEQUENCES: 1 | | | |
| : CORRESPONDENCE ADDRESS: | | | |
| : ADDRESSEE: Human Genome Sciences, Inc. | | | |
| : STREET: 9410 Key West Avenue | | | |
| : CITY: Rockville | | | |
| : STATE: MD | | | |
| : COUNTRY: USA | | | |
| : ZIP: 20850 | | | |
| : COMPUTER READABLE FORM: | | | |
| : MEDIUM TYPE: 3 1/2 inch diskette | | | |
| : COMPUTER: Dell Pentium | | | |
| : OPERATING SYSTEM: MS DOS v6.22 | | | |
| : SOFTWARE: ASCII Text | | | |
| : CURRENT APPLICATION DATA: | | | |
| : APPLICATION NUMBER: US/09/557,884 | | | |
| : FILING DATE: 25-Apr-2000 | | | |
| : CLASSIFICATION: <Unknown> | | | |
| : PRIOR APPLICATION DATA: | | | |
| : APPLICATION NUMBER: 08/476,102 | | | |
| : FILING DATE: JUN-5-1995 | | | |
| : ATTORNEY/AGENT INFORMATION: | | | |
| : NAME: Michelle S. Marks | | | |
| : REGISTRATION NUMBER: 41,971 | | | |
| : REFERENCE/DOCKET NUMBER: PB186P3 | | | |
| : TELECOMMUNICATION INFORMATION: | | | |
| : TELEPHONE: 301-309-8504 | | | |
| : TELEFAX: 301-309-8439 | | | |
| : INFORMATION FOR SEQ ID NO: 1: | | | |
| : SEQUENCE CHARACTERISTICS: | | | |
| : LENGTH: 1830121 base pairs | | | |
| : TYPE: nucleic acid | | | |
| : STRANDEDNESS: double | | | |
| : TOPOLOGY: linear | | | |
| : SEQUENCE DESCRIPTION: SEQ ID NO: 1: | | | |
| US-09-557-884-1 | | | |
| Query Match 23.0%; Score 664.4; DB 3; Length 1830121; | | | |
| Best Local Similarity 57.1%; Pred. No. 1.1e-159; | | | |
| Matches 1397; Conservative 1; Mismatches 1012; Indels 36; Gaps 9; | | | |
| Qy | 183 | AAATTTAACACAGAGTAAATAGGTGAACCTCACACCCAAAGAGCAGCAAGAAAAGTTCTTGT | 242 |
| Db | 564897 | ATAATAAAATTAAGGAATGAAAATATGCCTTAACCTTCAGAGAACCAAGAAAACTGATGC | 564838 |
| Qy | 243 | TATATTATCGGGCGAAGTGCTAGAAAGCCGAAAGCAGAGGGCTTAAGCTCAACCAAC | 302 |
| Db | 564837 | TTTTCTCGCGGGCGAACTTCGGCAAAACGCAAGCAGCGCGCTTAATTTAAACTATC | 564778 |
| Qy | 303 | CCGAAGCATTGCTTACATTAGTGGCCCATTTATGGAAGAGCGCGCTGGAAAAAAA | 362 |
| Db | 564777 | CAGAACTATTGCTTATATTCTAGTCAATTTACAGAGGCAGC--- | 564721 |
| Qy | 363 | CCGTTGCCAGCTTATGGAAGAGTGCACTTTTGAAGAAAGATGAAGTAATGCCCG | 422 |
| Db | 564720 | GTGTAGCGGAAGTCATGCAATATGCGCAACACACTTTTAAACCGTTGATGATGTCATGAAG | 564661 |

| | | | |
|----|--------|---|--------|
| Qy | 423 | GGGTGGGTAATATATGGTCTCCCATCTAGTGTAGAGGCCACCTTTTCTGTAGTGTAGCAAAAC | 482 |
| Db | 564650 | GTGTGGCGGAATGGTTTATGAAGTCCAGATTTGAAGCTACTTTTCCCGATGCGACGAAC | 564601 |
| Qy | 483 | TTGTAACTGTCAATTTGGCCCATC---CAACACAGATGAGCACTTCAAGCGCGCGAAGTG | 538 |
| Db | 564600 | TTGTACCGTGCATAATCCAATCAGATAACCGWAGGTGGGCTTTAGCCCAAAAATAAA | 564541 |
| Qy | 539 | AAATTTGGTTGCGATAAGACATCGAGCTCAATGCAGGCAAGAAGTAACCGAACTTGAG | 598 |
| Db | 564540 | AATATCAATGTGGGTAAAGCCACCCCTACAAGGAACAAAGATATGATCTCCAGCGCAAT | 564481 |
| Qy | 599 | GTTACTAATGAAGGCGCTAAATCTCTGC--ATGTGGGTAGCCATTTCCACTTCTTTTGAAG | 656 |
| Db | 564480 | ACCAATTAGCCGAAGGCGATATCTCGTAATGTGCGCAGAAAAACCGTAAAAATCGAAG | 564421 |
| Qy | 657 | CTAACAGGCACTAAATTTCCGATCGTGAAGAACCCCTATGGCAAAACGCTC-----AGATA | 710 |
| Db | 564420 | TAAACAATTTAGGCGACGCGCCCAATTTCAAGTTTGGCTCGCATTAACAATTTTGTGAACCA | 564361 |
| Qy | 711 | TTCCCTCTGGCAACACGCTACGCAATTTGGGGCAGGACAAACCCGCAAGGTGCAAGTTGATTC | 770 |
| Db | 564360 | ATAATGCCCTTAATTTGACCGCACTTTGGCACGTGGGCTTAAATGTTCCATCTG | 564301 |
| Qy | 771 | CTCTTGTGGCAGTAAAGAAAGTGAATGGCATGAACGGGCTTGTGAATAACATCGCGGATG | 830 |
| Db | 564300 | GCAATCGGTGCGTTTGAACCCGTGAAGTGAATCAGTGGAAATTGATGCTTTTGTGGT | 564241 |
| Qy | 831 | AACGCCATAACAAGCGCTTGACAGCGCGAATCTCACGGATTTATCAAGTAAGG-- | 888 |
| Db | 564240 | GTAACCA--AATCATTTTATGGTTTCCATAATCAAAATTGATGGCAAAATATAAGGTAGGCA | 564182 |
| Qy | 889 | AGACTTCCCATCAAAATGAAAAAACAAGAAATATGTAATACCTACGCGACCCACCAAGGCG | 948 |
| Db | 564181 | AGATGGCAATTAACAATTTCAAGAGCGCAATATAGCAACTTATGGTCCAAAGTTGGCG | 564122 |
| Qy | 949 | ATAAGTGCCTTAGGAGATACCGATCTTTGGGCAGAAAGTAGAAATGACTATACCACT | 1008 |
| Db | 564121 | ATAAGTCCGTTTAGGCGATACCAATTTATGGCAACCAATTGAAACAAGATTTATTGACCA | 564062 |
| Qy | 1009 | ATGGGAAGAAGCTTAATTTGGCGGGTAAAACTATCCGTGAGGATGAGGTGAGTCTAGAGC- | 1067 |
| Db | 564061 | AAGGTGATGAGTGTAAATTTTGGTGGCGGTAAAGCGTGCATGATGGTCTCAAGCG | 564002 |
| Qy | 1068 | -----AATAGCCCTGATGAAACACCCCTAGATTTTAGTTCATCACTAACCGGATGATTA | 1119 |
| Db | 564001 | GTACGGCAACTCGCGACAATCCGAATGTATTGGATTTTGTGATTAACAAGTATGATCA | 563942 |
| Qy | 1120 | TCGA---CTACACCGGGATTTACAAAGCCGACATTTGGGATTTAAAAACGCAAAATCCATG | 1176 |
| Db | 563941 | TTGATGCTAAATTAGGCATTTATCAAGCCGATATTGGTATTTCGTATGCGCGTATTGTGG | 563882 |
| Qy | 1177 | GCAATTGGCAAGGCAAGAAACAAGGACATGCAAGATGGCGTAAAGCCCTCATATGGTCTGG | 1236 |
| Db | 563881 | GTAATTGGACAGCAGGTAAACCTGCACACCATGGATAACGTCACAGCAAAATATGATTATCG | 563822 |
| Qy | 1237 | GTGTGGGCACAGACACTACGAGGAGGTATGATTATTACCCCTGGCGGAATCGATT | 1296 |
| Db | 563821 | GTGCAAGCAGCGAAGTTTATTAACCGGTGCACATTTAAATTGCAACCCGCTGGTGTATCGATA | 563762 |
| Qy | 1297 | CACACCCCACTTCTTCTCCACAAACAATTTCCCTACCGCTCTAGCCCAATGGCGTTACAA | 1356 |
| Db | 563761 | CCCACTTCACTTTATTGTTCACAAACAAGCACAACATGCAATTTGAAAGTGGCGTTACCA | 563702 |
| Qy | 1357 | CCATGTTTGGAGCGGCACAGGTCTGTAGATGGCAAGAAATGCGACTACTATCATCTCCGG | 1416 |
| Db | 563701 | CGTTAAATTGGTGGTGGAACTGGCCCTGTGTGTTACACAGCAACCACTTGTACCCCTG | 563642 |
| Qy | 1417 | GCAATTGAACCTGCACCGCATTTTGGCGCAGCAGAGAGATTTCTATGATGTGGCT | 1476 |
| Db | 563641 | GCGCATTGGTATATGGAACGTATGTTTCAAGCGCGCAAGAGCCCTTGCCTGTAAACGTCGGAT | 563582 |
| Qy | 1477 | TTTTTGGGCAAGGCAATAGTCTTAGCAAAAAACAATTTGTAGAAACAAGTAGAAGCGGGCG | 1536 |

| | | | | | | | |
|----|--------|---|--------|--------|---|---|--------|
| Qy | 423 | GGGTGGGTAAATATGTTCCCGATCTAGGTGTAGAGCCACCTTTCTGTGATGGTACGAAAC | 482 | 1477 | TTTTGGGCAAAAGGCAATAGCTCTAGCAAAAACAACCTTGTAGAACAAAGTAGAGCGGGC | 1536 | |
| Dd | 564660 | GTGTGGCGGAAATGGTTCATGAAGTCCAGATTGAAGCTACTTTCCCGGATGGCAGCAAC | 564601 | 563581 | TTTTTGTGTAAGGCAACCTGTTCAACCCCTAGATCTCTCGTGTGAGCAAAATGTAAGCGGGTG | 563522 | |
| Qy | 483 | TTGTACTGTGAATTTGCCCATC-----GAACCGATAGACACTTCAAGCGGGCGAAGTG | 538 | Qy | 1537 | CGATTGGTTTTAAATTTGCATGAAGACTGGGGCACAACCAAGTCGATCGATCACTGCT | 1596 |
| Dd | 564600 | TTGTTACCGTGCATAATCCCAATCAGATAACCGWGGGTGGCTTTAGCCCAACAAATATAA | 564541 | Dd | 563521 | CATTAGGTTTTAAANAATCCAGAGACTGGGGTGCAACGCTCGCGTGATTTGATTCTGCCT | 563462 |
| Qy | 539 | AAATTTGGTTCGATGAAGACATCGAGCTCAATGTCAGGCAAAAGTAACCGAACTTTGAG | 598 | Qy | 1597 | TGAGCGTGGCAGATGAATPACGATGTGCAAGTTTGTATCCACACCGATACAGTCAATGAGG | 1656 |
| Dd | 564540 | AATATCAATGGTGGCTTAAAGCCACCTACAGGAACAAGATATGATCCAGCGCAAT | 564481 | Dd | 563461 | TAAAGTAGCAGATGAATGGATATTCAAGTGGCCATTACACAGACACGCTAAATGAAA | 563402 |
| Qy | 599 | GTTTACTAATGAAGGGCTTAATCTTGC--ATGTGGGTAGCCATTTCCACTTCTTTGAAG | 656 | Qy | 1657 | CAGGTTATGTAGATGACACCCCTAAATGCAATGAACGGGGCGCCATCCATGCTACCA | 1716 |
| Dd | 564480 | ACCAATTAGCGGAAGCGATTTCTCGCTAATGTGCGCAGAAAAACCGTAAAAATCGAAG | 564421 | Dd | 563401 | GTGGCTTTTGGAAAGACGATGAAGCGATTGATGAGCGAGTCAATTCATCTTTCCATA | 563342 |
| Qy | 657 | CTAACAGGCACATAAATTCGATCGTGAAAAAGCTATGGCAAAAGCCT-----AGATA | 710 | Qy | 1717 | TTGAGGAGCGGGTGGAGGACACTCACCTGATGTTATCACCATGCGAGCGAGCTCAATA | 1776 |
| Dd | 564420 | TAACAAAATTCAGCGACCGCCCAATTCAGTTGGCTCGCATTACCATTTTTTTGAAACCA | 564361 | Dd | 563341 | CGGAGGGCGCAGGTGGTGGTCATGCACCTGACATCATTAAGCAGCGATGTATTCAAACG | 563282 |
| Qy | 711 | TTCCCTCTGGCAACGCTACGATTTGGGCGAGCAAAACCGCAAGTGCAGTTGATTC | 770 | Qy | 1777 | TTCTACCTCTCTCCACCAACCCCTCACTATTCCTATACCATTAAATACGGTTGCAGAA | 1836 |
| Dd | 564360 | ATAATGCCCTTAAATTTGACCGCACTTTGGCACGTGGAATGCGCTTAAATGTTCCTACTG | 564301 | Dd | 563281 | TATTACCTGCTTCAACCAACCCGACTCGTCCGTTTACCAAAACACCATTTGATGAACAT | 563222 |
| Qy | 771 | CTCTTGGTGGCAGTAAAAAAGTGAATGGCATGAACGGGCTTGTGAATAAACATCGCGGATG | 830 | Qy | 1837 | TAGACATGCTCATGACATGCCACCACTAGACAAACGATCCGCGAGGATTTTACAATTTT | 1896 |
| Dd | 564300 | GCNAATGGGTGGCTTTTGAACCGGCTGAAGTGAATCAGTGGAAATAGTTGCTTTTGGTG | 564241 | Dd | 563221 | TGGATATGTTGATGGTTTGCCATCACTTAGATAAAACGGCTGCCGGAAGACGTAGCTTTTG | 563162 |
| Qy | 831 | AACGCCATAAACATAAAGCGCTTGACAAGCGGAAATCTCAGGATTTTATCAAGTAAG-- | 888 | Qy | 1897 | CTCAAGCGGTATCGCCCCCGCTTATCGCGGCTGAGATGTGCTCCATGATATCGGGTG | 1956 |
| Dd | 564240 | GTAACCA-AATCATTTATGTTTCCATAATCAAAATGATGGCAAAATTAAGGTAGGGCA | 564182 | Dd | 563161 | CCGATAGCGGTATCGGCCCTGAAACCAATTGAGCAGAGAATATTTTGATGATATGCGGCG | 563102 |
| Qy | 889 | AGACTCCCATGAATAAAGAAACAGAAATATGTAATACCTACGACCCACCAAGGCG | 948 | Qy | 1957 | TGATCGGATCACAAAGCTCGGATTCGCAAGCAATGGGGCTGCAGCGAAGTGAATTCCTC | 2016 |
| Dd | 564181 | AGATGGCATTAACAAATTTCAAGAGCGCAATATGTAGCAACTTATGCTCCAACAGTTGGCG | 564122 | Dd | 563101 | TCCTTCTCAATTATGAGTTTCACTCTCAAGCGATGGAGATTTTGGCGAAGTCGTTATTC | 563042 |
| Qy | 949 | ATAAAGTGGCTTAGAGATACCGATCTTTGGCGAGAAATAGAACATGACTATACCACT | 1008 | Qy | 2017 | GAACTTGGCAGACTCGCGATAAGAAATAAAAGAAATTTGGTAAGCTTCTCGAAGATGGCA | 2076 |
| Dd | 564121 | ATAAAGTCGTTTAGCGATACCAATTTATGGCAACCATTTGAACAAGATTTATGACCA | 564062 | Dd | 563041 | GTACATGCAAACTGCAGATAAGATAAGATAAGTGCACGTGGTGAAGTAAAGGA-562983 | |
| Qy | 1009 | ATGGCGAAGAACTTAAATTTGGCGCGGTAAAACTATCCGTAGGGGTATGGTCAGAGC- | 1067 | Qy | 2077 | AAGATAAGATAATTTCCGCAATTAAGCGCTACATCTCCAAATACACTATCAACCCCGCTT | 2136 |
| Dd | 564061 | AAGGTATGATGTAAATTTGTGGGGGTAAAGCGGTGCGTGTGATGGCTCAGAGCG | 564002 | Dd | 562982 | -----AACGATAACTTCCGTTATTAACGATATATCGGAAATACACCATCAACCCAGCAA | 562928 |
| Qy | 1068 | -----AATAGCCCTGATGAATAACCCCTAGATTTAGTATCATCACTAACCGATGATTA | 1119 | Qy | 2137 | TGACCCACGGCGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCAAGATCGCGACTTGG | 2196 |
| Dd | 564001 | GTACGGCAACTCGCGCAATCCGAATGTATTGGATTTTGTGATTTACCAACGTGATGATCA | 563942 | Dd | 562927 | TTGCACATGGTATTGCGGAGCATATTGGCTCGTTAGAAGTGGGTAAATTCGAGATATCG | 562868 |
| Qy | 1120 | TGCA---CTACACGGGATTTACAAAGCGACATTTGGGATTAATAACCGCAAAATCCATG | 1176 | Qy | 2197 | TGGTGTGAAATCTCGCCTTTTTTGGCGTAAACCCCAAAATCGTGATCAAAAGGCGGTATGG | 2256 |
| Dd | 563941 | TTGATGCTAAATTAGGCAITTTCAAGGCCATATTGATATTCGTGATGGGCGTATTGTGG | 563882 | Dd | 562867 | TGTTATGGAACCGATGTTCTTTGGCGTAAACCTGNAAGTGGTGAATAAAAAGGCTTTA | 562808 |
| Qy | 1177 | GCATTGGCAAGCAGGAAACAAGACATGCAAGATGGCGTAAAGCCCTCATATGTCGTGG | 1236 | Qy | 2257 | TGGTCTTCTCGAAATGGCGATTTCTAACCGGTCTGTGCCCACTCTCCCAACCGGTTTAT | 2316 |
| Dd | 563881 | GTAATTGCAACAGCAGGTAAACCTTGACACCATGATTAAGTCAACCAAAATGATATCG | 563822 | Dd | 562807 | TTAGCTATGCGAAAATGGCGATCCAAATGCCTCAATTTCCAAACCGCAACCTGTATTCT | 562748 |
| Qy | 1237 | GTGTGGGCACAGAACCACTAGCAGGGGAAGGTATGATTTATTCGCTGGGGGAATCGATT | 1296 | Qy | 2317 | ACCGGAAATGTTTGGGCATCACGGCAGGCGAAATTTTGACACCGAGCATCACTTTTGT | 2376 |
| Dd | 563821 | GTGCAAGCACGGAAGTTCATAACCGGTGCACATTTAATTTGCAACCGCTGGTATCGATA | 563762 | Dd | 562747 | ACCGTCCAATGTACGGTGCACAAGGCTTTAGCAACCGCAACCAACGACGATTTCTTTG | 562688 |
| Qy | 1297 | CACACACCACTTCTCTTCTCCACCAACATTCCTACCGCTCTAGCCAAATGGCGTTACAA | 1356 | Qy | 2377 | CCAAAGTCGCTTATGAAATGGGTGAAGAAAAGCTGGGCTTTAGAGCGCCAAAGTCTAC | 2436 |
| Dd | 563761 | CCACATCTCACTTTATTTGTGTCACAAACAGCAACATGCAATTTGAAAGTGGCGTTACCA | 563702 | Dd | 562687 | CACAAGCGCTGAAAAGCTGATATTCTGTGGAAGTTCGTTTACAAAGAAACCATTTG | 562628 |
| Qy | 1357 | CCATGTTTGGAGCGCACAGGTCCTGTAGATGCGCAGATGCGACTACTATCACTCCGG | 1416 | Qy | 2437 | CGGTCAAAAACCTGCGTAACATCAACAAAGAGCTTCAAGTTTCAACGCAAAACCGCAA | 2496 |
| Dd | 563701 | CGTTAATTTGGTGGAACTGGCCCTGCTGATGGTACACAGCAACCACTTTGACCCCTG | 563642 | Dd | 562627 | CTGTGAAAGGCTCGCGCAACGCTAGTAAAAAAGATCTGGTTCATAATGATGTAACACCAA | 562568 |
| Qy | 1417 | GCAAAATGGAATTTGCACCGCATTTGCGCGCAGCAGGAAGATTTCTTATGAATGTGGGCT | 1476 | Qy | 2497 | AAATCACCGTCGATCCGAAAAACCTTCGAGGCTTTGTAGATGGCAAACTCTGCACCTCTA | 2556 |
| Dd | 563641 | GGCATGGTATATGGAACGATATGTTTCAAGCGCGAAGACCTTTCGCGGTAACCGTCGGAT | 563582 | Dd | 562567 | ACATTACTGTTGATGCTGAACGTTTATGAAGTTTCGATGGAGCGAGAGATTAATACCTGTG | 562508 |
| | | | | Qy | 2557 | AACCACTCTCGCAAGTGCCTCTAGCCCGACGCTACACTTCTTCTTA | 2602 |

LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature

Query Match 23.0%; Score 664.4; DB 3; Length 1830121;

Best Local Similarity 57.1%; Pred. No. 1.1e-159;

Matches 1397; Conservative 1; Mismatches 1012; Indels 36; Gaps 9;

QY 183 AAATTAAACAGAGTAGTAGGTGAAACTCACACCCCAAGAGCAAGAAAGTTCTTGT 242

DB 564897 ATAATAAATTAGGAATGAAAATATGCATTAACTTCCAGAGCAAGAAACTGATGC 564838

QY 243 TATATTATCGGGCGAAGTGGGTAGAAAGCGCAAGCAGAGGGCTTTAAAGCTCAACCAAC 302

DB 564837 TTTTCTCGGGCGGAGTCTGGCGCAACAGCAAGCAGCGGCTTAATTAACCTATC 564778

QY 303 CCGAAGCCATTGCTTACATTAGTCCCATATTATTGGACGAAGCGCGCTGGAAAAAAA 362
DB 564777 CAGAAAACATTGCTTATATTGCTAGTCAATTTTACAAGAGGCAGC---AAGAGAAGGAATGA 564721
QY 363 CCGTTGCCAGCTTATGGAAGAGTGCATGCACATTTTGGAAAAAAGATCAAGTAAGTCCCG 422
DB 564720 GTGTAGCGGAAGTCATGCAATATGCGCGAACACATTTTAAACCGTTGATGATGTCTAGGAAG 564661
QY 423 GGGTGGGTAATATGTTTCCCGATCTAGGTGTAGAACCCACCTTTCTCTGATGTTAGAAAC 482
DB 564660 GTGTGGCGGAATGTTTCATGAAGTCCAGATTGAAGCTACTTTCCCGCATGSCACGAAC 564601
QY 483 TTGTAAGTGTGAATGTCGCCATC---GAACCAAGATGAGCACTTCAAAAGCGGCGCAAGTG 538
DB 564600 TTGTTACCGTCATAATCCAATCAGATAACCGWAGGGTGGGCTTTAGCCCAAAAATAAA 564541
QY 539 AAATTTGGTTCGATAAAGACATCAGACTCAATCAGCGCAAGAGTAACCGAACTTGAG 598
DB 564540 AATATCAATGGTGGCTTAAAGCCACCTTACAAGAAACAAAGATATGATCCCGGCGAAT 564481
QY 599 GTTACTAATGAAGGGCTTAAATCCTTGC--ATGTGGGTAGCCATTTCCACTTTCTTTGAAG 656
DB 564480 ACCAATTAGCGGAAGCGATATTTCTGCTAATGTGCGCAGAAAAACCGTAAAAATCGAAG 564421
QY 657 CTAAACAGGCATAAAATTCGATCGTGAAGAAAGCCTATGGCAACGCCT-----AGATA 710
DB 564420 TAACAAATTCAGGCGACCGCCCAATTCAGTTGGCTCGCAATTACCATTTTTTTGAAACCA 564361
QY 711 TTCCCTCTGGCAACACGCTAGCATTGGGGCAGACAAACCCGCAAAAGTCAGTTGATTC 770
DB 564360 ATAATGCCCTTAAATTTGACCGCACTTTGGCAGTCGTAATGCGTTAAATGTTCCATCTG 564301
QY 771 CTCTTGGTGGCAGTAAAAAAGTGTGSCATGAAAGGCTTTGTGAATAACATCGCGGATG 830
DB 564300 GCAATGCGGTGCGTTTTGAAACCCGCTGAAGTGAATCAGTGAATTAGTGTCTTTTGGTG 564241
QY 831 AAGCCATAAACAATAAGCGCTTGACAAGGCGAAATCTCACGGATTTATCAAGTAAGG-- 888
DB 564240 GTAACCA-NATCAITTTATGTTTCCATAATCAAATTTGATGCGCAAAATTAAGGTAGGCA 564182
QY 889 AGACTCCCATGAAATGAAAAAACAAGAAATATGTAATATCTACGGACCCCAACCAAGGG 948
DB 564181 AGATGGCATTAACAATTTCAAGAGCGCAATATATGCAAACTTATGCTTCAACAGTTGGCG 564122
QY 949 ATAAAGTGGCTTAGGAGATACCGATCTTTGGGCAGAAAGTAGAATCACTATACCACT 1008
DB 564121 ATAAAGTCCGTTTTAGCGGATACCAATTTATGGGCAACCAITTTGAACAAGATTTATTGACCA 564062
QY 1009 ATGCGGAAGAACTTAAATTTGGCGCGGTAAAACTATCCGTGAGGTATGGTCAGAGC- 1067
DB 564061 AAGTGATGATGTAAATTTGGTGGCGGTAAAGCGTGGTGATGTTAGTCTCAAGCG 564002
QY 1068 -----AATAGCCCTGATGAAAAACCCCTAGATTTTAGTCATCATCAACCGGATGATTA 1119
DB 564001 GTACGGCAACTCGCGCAATCCGAATGTATTGGATTTGTGATTACCAACCGTATGATCA 563942
QY 1120 TCGA---CTACACGGGATTTAAACCGGCATTTGGATTTAAAAACCGCAAAATCCATG 1176
DB 563941 TTGATGCTAAATTTAGGCAITTTCAAAGCCGATATTGGTATTTCGTGATGGGCGTATTGTGG 563882
QY 1177 GCATTGCGAGGCAAGGAAAAACAAGACATGCAAGATGCGTAAGCCCTCATATGGTCGTGG 1236
DB 563881 GTATTGCAAGCAGGTAAACCTTGACACCTAGGATAGTCAACCAATATGATTTATCG 563822
QY 1237 GTGTGGCGCAGAAAGCACTAGCAGGGGAAGGTATGATTTATTACCGCTGGGGGAATTCGATT 1296
DB 563821 GTCAAGCAGGAAGTTCATAACCGTGCACATTTAAATTGCAACCGCTGCTGTTATCGATA 563762
QY 1297 CACACACCACTTCTCTTTTCCACAAACAATTTCCCTACCGCTCTAGCCAAATGGCGTTACAA 1356
DB 563761 CCCACATTCATTTTATTGTTCCCAACAACAGCAACATGCAATTTGAAATGGCGGCTTACCA 563702
QY 1357 CCATGTTTGAGGGCGCACAGGTCTCTGTAGATGSCACGAATGCGCACTACTATCACTCCGG 1416

QY 190 ACACAAGGAGTAATAGGTGAAACTCACACCAAGAGCAAGAAAAGTTCTTGTTATATTA 249
Db 4028 AATTGGGGGTTTTAAATTTGCAATTTTACACAACGAGAGCAAGACAAATTAATGTTAGT 3969
QY 250 TCGCGCGAGTGGCTAGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCCGAAGC 309
Db 3968 GCGCGCGAGTGGCTAGCAAGTGGCTGTAAGACGCTGTTTGAATACTAAATCATCTGAGGC 3909
QY 310 CATTTGCTTACATTTAGTCCCATATTTATGACGAAGCGCGCTGGAAGAAAAAACCCTTGC 369
Db 3908 ATTAGCTTTAATCAGCGATGAATATTAGAAAGTGCACGCGATGG---TAAGACCGTTC 3852
QY 370 CCAGCTTATGGAAGAGTGCATGCATTTTGAAGAAAAGATGAAGTAATGCCCCGGGTGG 429
Db 3851 AGAGTTAAATGAGTTATGGTATAGACAAATTTCTAAACAAAGAGAGTGTATGATGGTGTGCA 3792
QY 430 TAATATGGTTCCCGATCTAGGTGTAGNAGCCACTTTCTTGATGGTACGAACTTGTAAAC 489
Db 3791 ACACATGATTTACAGATATCGAAATCGAGGCTACGTTCCCCGATGGTACTAAGTTAATCAC 3732
QY 490 TGTGAATTTGGCCCATCGAACCATAGTGAAGCACTTTCAAAGCGGGC- GAAGTGAATAATTTGGTT 548
Db 3731 AGTACATCACCTTATTTTAAAGGAGGAGTCAATATGATACAGAGAGAAATTTATTACAA 3672
QY 549 GCGATAAAGACATCGAGCTCAATTCGAGGCAAGAAAGTAAACCGAATTTAGGTTACTAATG 608
Db 3671 AAAGTACAGAGGTTGAAATTTAATAACCATCATCTCGAATACAGTTATCGAAGTTGMAAATA 3612
QY 609 AAGGGCTTAAATCTTTCATGTGGGTAGCCATTTCCATTTCTTTTGAAGCTTAAACAGGCAC 668
Db 3611 CAGGAGATCGACCAATTTCAAGTGGGTGTCATTTTCAATTTTATGAAGCAAAATGCGAGCAT 3552
QY 669 TAAAAATTCGATCGTGAAGAAAGCTATGGCAAGCCCTAGATATTTCCCTCTGGCAACACGC 728
Db 3551 TAGATTTGAAACGTGAAATGGCATATGGAAAAATTTAGATATTCAGCTGGAGAGCTG 3492
QY 729 TAGCCATTTGGGCGAGACAAACCGCAAGTGCAGTTGATTCCTCTTGGTGGCAGTAAAA 788
Db 3491 TTCGATTTGAACTGGGGATAAAAAGAGTTCAAATTTAGTTGAATGCTGGCNAACGTA 3432
QY 789 AAGTGAATGGCATGAACGGGCTTTGTGAA---TAAACATCGCGGATGAACGCCATAAACAATA 845
Db 3431 AAATTTTGTGTTTCGTGGTATGGTCAATGGTCTATCGATGAGTCACGTGTCTATCGCC 3372
QY 846 AAGCGCTTGACAGGGCAATCTCAAGGATTAATC----- 880
Db 3371 CAACATGATAAATGATGAATATGCAAGGTGATTCGGAGATAACGGTGTGAAACAGTGA 3312
QY 881 -----AAGTAAGGAGACTCCCATGAAATGAAATAAAGATAATGATAGAGCTT 928
Db 3311 ATAAAAAAGGAGGAAAAGATCATGAGCTTTAAATGACCGCAAAATCAATATAGGAGCTT 3252
QY 929 CTACGACCCCAACAAAGGCGATAAGTGGCTTTAGGAGATPACCGATCTTTTGGCGAGAAGT 988
Db 3251 ATACGGTCAACCTGTTGGAGATTCATTCGTTTAGTGATACGAATCTATTGCTCAAT 3192
QY 989 AGAACATGACTATACCACTATGGCGAAGAACTTAAATTTGGCGCGGTTAAACCTATCCG 1048
Db 3191 AGAAAAAGACTATGCGGTTTATGGTGAAGAAAGTACTTTTGGTGGTGTAAATCTATTAG 3132
QY 1049 TGAGGCTATGGGTACAGCAATAGCCC-----TGATGAAAACACCCCTAGATTAGT 1099
Db 3131 AGACGGTATGGCGCAAAATCTCTGTGTAAACGCTGATGAGTGAACGCTTCAGACCTTGT 3072
QY 1100 CATCACTAAACGCGATGATTATCGACTACACCGGATTTTACAAAGCGCAATTTGGGATTA 1159
Db 3071 CATTTCTAATGCGTTATTATCGATTACGATAAGTGGTTTAAAGCTGATATAGCAATTAA 3012
QY 1160 AAACGCAAAATCCATGGCATTTGGCAAGGAGGAGAAACAAAGCAATGCAAGATGCGGTAG 1219
Db 3011 AAATGGTTATATTTTCGCCATAGGTAAATGCGGCAACCCAGATATAATGATAATGT--- 2955
QY 1220 CCTCATATGGTGTGGGTGGGCACAGNAGCACTAGCAGGGGAAGGTATGATTATAC 1279

Db 2954 ---CGACATTATTATAGTTCAACAACAGATATCATTCGCGTGAAGGTAAATGTCAC 2898
QY 1280 CGCTGGGCGAATCGAATTACACACCCACTTCCTTTTCCCAACAATAATCCCTACCGCTCT 1339
Db 2897 TGTGGTGGTATTGATACTCATGTTTATTATTAATCTGAACAAGCAGAGGTGCGATT 2838
QY 1340 AGCCAATGGGCTTACAACCATGTTTGGAGGCGGCACAGGTCTGTAGATGGCAGAAATGC 1399
Db 2837 AGAAAGTGGTATTACGACTCATATTGGTGGTGTACTGCTGCTTCTGAGAGGTCTTAAAGC 2778
QY 1400 GACTACTATCACTCCGCGCAAAATCGAACTTGCACCGCATGTTGCGCGAGCAGAGAGATA 1459
Db 2777 AACAACTGTAACTCCAGGTCCATGGCATATTTCATAGAAATGTTAGAAAGTGTCCGAAAGTTT 2718
QY 1460 TTCTATCAATGTGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAACCTTGTAGA 1519
Db 2717 ACCGATTAAATGTCGTTTTTACAGGTAAAGCAAGCAACAATCCAACTGCACTCATTTGA 2658
QY 1520 ACAAGTAGAAGCGGCGCGAATTTGGTTTAAATTTCAATGCAAGACTGGGGCAACAACCAAG 1579
Db 2657 ACAATCAATCCGAGCAATTTGGATTAAAGATACATGAAGACTGGGGTGCACCAACCATC 2598
QY 1580 TGGCATCGATCACTGCTTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACAC 1639
Db 2597 TGTCTTGAGTCATGCATTTAGATGTTGCTGTGATGAAATTTGATGTTCAAAATTTGCAATTACATGC 2538
QY 1640 CGATACAGTCAATGAGCGAGGTTATGTAGATGACACCCCTAAATGCAATGAACGGCGCGC 1699
Db 2537 AGATACTTTTAAATGAAGCAGGATTTATGGAAGACAAATGGCTGCTGTTTAAAGACGCTGT 2478
QY 1700 CATCCATCGCTTACCACATTTGAGGAGGGGTGGAGGACACTCACCTGATGTTTATCACCAT 1759
Db 2477 ACTTCATATGACATCTAGCTGAAGGTGCTGGTGGCGGTCTATGGCCTGATTTAATTTAAATC 2418
QY 1760 GGCAGGCGAGCTCAATATTTCTACCTCTCCACCAACCCCACTATPTCCCTATATACCAATTA 1819
Db 2417 CGCTGCAATTTTCAAAATATTTTACCTTCATCTACAAATCCAACTTTTGCCTTATACACATA 2358
QY 1820 TACGTTTGCAGAACACTTTAGACATGCTCATGACATGCCACCCTAGACAAACGGATCCG 1879
Db 2357 TACTGTAGATGAAACATTTAGATATGGTAAATGATTACTCACCATTTTAAATTCGGGCTATTCC 2298
QY 1880 CGAGGATTTTCAATTTTCTCAAAGCGGTATCCGCGCCCGCTCTATCGCGGCTGAAGATGT 1939
Db 2297 TGAAGATATCGCATTCGAGATTCCAGTATTCGTAAAGAAACGATTCGAGCAGAGATGT 2238
QY 1940 GCTCCATGATATGGGTGTGATCGCGATCAAGCTCGGATTCGCAAGCAATGGGGCGTGC 1999
Db 2237 TCTGCAAGATATGGGTGTATTCAGTATGATTTAGTTCCGATTCACAAGCAATGGGCGGTGT 2178
QY 2000 AGCGAAGTATTCCTCGAACTTGGCAGATCGCGATGAAGATAAATAAGAAATTTGGTAA 2059
Db 2177 AGGTGAAGTAAATTACACGAAACATGGCAATGAGACATCGCATGAAGAAACCAACCTGGTCC 2118
QY 2060 GCTTCCTGAAGATGGCAAGATAACGATAATTTCCGCAATTAAGCGCTACATCTCCAAATA 2119
Db 2117 TTTAGATGGTGAATTTTGAACATAATGATAATAATCGCATCAACGTTTATATCGTAAATA 2058
QY 2120 CACTATCAACCCCGCTTTGACCCACGGCGTGGAGGAGTATATCGGCTCTGTGGAAAGAGG 2179
Db 2057 TACATTTAACCCAGCAATTTACATGTTATTTCTGAAATATGTAGGATCTATCGAGCCGG 1998
QY 2180 CAAGATCCGACTTGGTGGTGGAAATCTCGCCCTTTTGGCGTAAACCCCAAAATCGT 2239
Db 1997 CAAACTAGCTGACATTTGCTTATGGGACCCCAATTTTCTTTGGGGTTAAACCTGGAATAGT 1938
QY 2240 GATCAAGGCGGATGTTGGTGTCTCTGAAATGGCGATTTCTAACGGCTGTGCGCCAC 2299
Db 1937 TGTAAGGGCGGATTAATACTCTGCGGTAAATGGCGATCAAAATGGTTCTTATACCTAC 1878
QY 2300 TCCCAACCGGTTTATTACCGCAAAATGTTTGGGCAATCAGCGCAAGGCGGAAATTTGACAC 2359

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 28, 2005, 08:00:02 ; Search time 14.5164 Seconds
(without alignments)
1497.963 Million cell updates/sec

Title: US-09-904-994B-2
Perfect score: 1181
Sequence: 1 VKLTPKEQKFLYYAGEVA.....ADRRHKHAKDKAKSHGFIK 226

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80:*
1: Piri:*
2: Piri2:*
3: Piri3:*
4: Piri4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 613.5 | 51.9 | 238 | 2 B71977 | urease alpha chain |
| 2 | 610.5 | 51.7 | 238 | 1 URKCAP | urease EC 3.5.1.5 |
| 3 | 566.5 | 48.0 | 237 | 1 S35290 | urease EC 3.5.1.5 |
| 4 | 483 | 40.9 | 228 | 2 D75586 | urease, beta/gamma |
| 5 | 439.5 | 37.2 | 840 | 1 URJB | urease EC 3.5.1.5 |
| 6 | 419.5 | 35.5 | 838 | 2 A96599 | probable urease F1 |
| 7 | 398.5 | 33.7 | 835 | 2 T37939 | urease EC 3.5.1.5 |
| 8 | 395 | 33.4 | 227 | 2 T29055 | urease EC 3.5.1.5 |
| 9 | 281.5 | 23.6 | 100 | 1 A36138 | urease EC 3.5.1.5 |
| 10 | 278.5 | 23.6 | 100 | 2 H70664 | probable ureA prot |
| 11 | 277.5 | 23.5 | 100 | 2 D85603 | probable urease st |
| 12 | 277.5 | 23.5 | 100 | 2 B90794 | urease gamma subun |
| 13 | 276.5 | 23.4 | 100 | 2 S74889 | urease EC 3.5.1.5 |
| 14 | 271 | 22.9 | 105 | 2 B69729 | urease gamma subu |
| 15 | 270.5 | 22.9 | 100 | 2 A36950 | urease EC 3.5.1.5 |
| 16 | 269.5 | 22.8 | 100 | 2 S47102 | urease EC 3.5.1.5 |
| 17 | 267.5 | 22.7 | 100 | 2 A64076 | urease EC 3.5.1.5 |
| 18 | 266 | 22.5 | 137 | 1 S38484 | urease EC 3.5.1.5 |
| 19 | 265.5 | 22.5 | 100 | 2 C90027 | urease gamma subun |
| 20 | 264.5 | 22.4 | 100 | 2 B47090 | urease EC 3.5.1.5 |
| 21 | 263 | 22.3 | 124 | 2 C69729 | urease (beta subun |
| 22 | 262.5 | 22.2 | 100 | 1 S08478 | urease EC 3.5.1.5 |
| 23 | 261.5 | 22.1 | 100 | 1 B43719 | urease EC 3.5.1.5 |
| 24 | 259.5 | 22.0 | 100 | 2 AD2264 | urease gamma chain |
| 25 | 258.5 | 21.9 | 100 | 2 E83037 | urease gamma subun |
| 26 | 257.5 | 21.8 | 100 | 2 S42502 | urease EC 3.5.1.5 |
| 27 | 256.5 | 21.7 | 100 | 2 D83681 | urease gamma subun |
| 28 | 251.5 | 21.3 | 100 | 2 AE2872 | urease gamma subun |
| 29 | 251.5 | 21.3 | 107 | 2 F97448 | urease gamma chain |

| | | | | | |
|----|-------|------|-----|----------|--------------------|
| 30 | 247 | 20.9 | 108 | 2 AE2264 | urease beta chain |
| 31 | 246.5 | 20.9 | 100 | 2 S38483 | urease EC 3.5.1.5 |
| 32 | 245 | 20.7 | 136 | 2 D90027 | urease beta subun1 |
| 33 | 242.5 | 20.5 | 100 | 2 A10324 | urease EC 3.5.1.5 |
| 34 | 241.5 | 20.4 | 164 | 1 S31418 | urease EC 3.5.1.5 |
| 35 | 239.5 | 20.3 | 126 | 2 S47103 | urease EC 3.5.1.5 |
| 36 | 237.5 | 20.1 | 101 | 2 AC2872 | urease beta subun1 |
| 37 | 236 | 20.0 | 101 | 1 S42604 | urease EC 3.5.1.5 |
| 38 | 234 | 19.8 | 105 | 2 S76224 | hypothetical prote |
| 39 | 232.5 | 19.7 | 100 | 2 AC3333 | urease EC 3.5.1.5 |
| 40 | 232.5 | 19.7 | 102 | 2 S31417 | urease EC 3.5.1.5 |
| 41 | 230.5 | 19.5 | 101 | 2 A82891 | urease complex com |
| 42 | 230 | 19.5 | 159 | 2 AB3333 | urease EC 3.5.1.5 |
| 43 | 228.5 | 19.3 | 100 | 2 T50708 | urease EC 3.5.1.5 |
| 44 | 225.5 | 19.1 | 101 | 1 I64075 | urease EC 3.5.1.5 |
| 45 | 225.5 | 19.1 | 101 | 2 AG3458 | urease EC 3.5.1.5 |

ALIGNMENTS

RESULT 1

B71977 urease alpha chain - Helicobacter pylori (strain J99)

C:Species: Helicobacter pylori

A:Variety: strain J99

C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004

C:Accession: B71977

R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;

; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;

Nature 397, 176-180, 1999

A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path

A:Reference number: A71800; MUID:99120557; PMID:9923682

A:Accession: B71977

A>Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-238 <ARN>

A:Cross-references: UNIPROT:Q9ZM24; UNIPARC:UPI0000137D6F; GB:AE001439; NII

A:Experimental source: strain J99

C:Genetics:

A:Gene: ureA

C:Superfamily: urease, fused gamma/beta subunit; urease 11K chain homology; urease 12K c

F:1-101/Domain: urease 11K chain homology <U11>

F:108-204/Domain: urease 12K chain homology <U12>

Query Match 51.9%; Score 613.5; DB 2; Length 238;

Best Local Similarity 53.1%; Pred. No. 1.3e-43;

Matches 119; Conservative 42; Mismatches 62; Indels 1; Gaps 1;

Qy 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAIVISAHIMDEARRGKKTVAOLME 60

Db 1 MKLTPKEDLMLHYAGELARRKKEGKIKLNTVEAVALISAHIMEARRGKKTVAOLME 60

Qy 61 CMHFKKDEVPVGNMVPDLGVEATPDGTLKVTNMPIDEPDFHFKAGEVFGCKDIE 120

Db 61 GRTLLKPDVWDGVASMIHEVGIEAMFDGTLKVTNMPIDEPDFHFKAGEVFGCKDIE 119

Qy 121 LNAGKEVTELVTEGPKSLHGVSHFFFEANKALKFKREKAYKRLDIPSGNTLRIGAG 180

Db 120 INEKKAVSVKVKVGNDRPVQIGSHFFFEVNRCLDFDREKTFGKRLDIASGTAVRFE 179

Qy 181 QTRKVLQIPGGSKKVIKMGVLNNIADRRHKHAKDKAKSHGF 224

Db 180 EKSVELIDIGNRRIFGNALVDRQADNESKIALHRAKERGF 223

RESULT 2

URKCAP

urease (EC 3.5.1.5) 26K chain - Helicobacter pylori (strains 26695 and others)

N:Alternate names: urea amidohydrolase; urease alpha chain

C:Species: Helicobacter pylori

C:Date: 30-Sep-1991 #sequence_revision 15-Aug-1997 #text_change 09-Jul-2004

C:Accession: A38537; A64529; S07884; B35306; A41502; A61371; E49215

R;Labigne, A.; Cussac, V.; Courcoux, P.
J. Bacteriol. 173, 1920-1931, 1991
A;Title: Shuttle cloning and nucleotide sequences of Helicobacter pylori genes responsible for urease
A;Reference number: A38537; MUID:91161505; PMID:2001995
A;Accession: A38537
A;Molecule type: DNA
A;Residues: 1-238 <LAB>
A;Cross-references: UNIPROT:P14916; UNIPARC:UPI000002C3CD; GB:M60398; NID:g149007; PIDN:
R;Tom, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.A.; Khalak, H.G.; Glodek, A.; McKenna
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.
Nature 388, 539-547, 1997
A;Authors: Wallin, S.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A;Reference number: A64520; MUID:97394467; PMID:9252185
A;Accession: A64529
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-238 <TOM>
A;Cross-references: UNIPARC:UPI000002C3CD; GB:AE000529; GB:AE000511; NID:g2313152; PIDN:
A;Experimental source: strain 26695
R;Clayton, C.L.; Pallen, M.J.; Kleanthous, H.; Wren, B.W.; Tabagchali, S.
Nucleic Acids Res. 18, 362, 1990
A;Title: Nucleotide sequence of two genes from Helicobacter pylori encoding for urease
A;Reference number: S07884; MUID:90221820; PMID:2326167
A;Accession: S07884
A;Molecule type: DNA
A;Residues: 1-36, 'R', 38-48, 'R', 50-131, 'PP', 134-238 <CLA>
A;Cross-references: UNIPARC:UPI000002DB17; EMBL:X17079; NID:g43633; PIDN:CAA34932.1; PID
R;Dunn, B.E.; Campbell, G.P.; Perez-Perez, G.I.; Blaser, M.J.
J. Biol. Chem. 265, 9464-9469, 1990
A;Title: Purification and characterization of urease from Helicobacter pylori.
A;Reference number: A35306; MUID:90264448; PMID:2188975
A;Accession: B35306
A;Molecule type: protein
A;Residues: 1-20 <DUN>
A;Cross-references: UNIPARC:UPI0000030924
R;Hu, L.T.; Mobley, H.L.T.
Infect. Immun. 58, 992-998, 1990
A;Title: Purification and N-terminal analysis of urease from Helicobacter pylori.
A;Reference number: A41502; MUID:90202165; PMID:2318539
A;Accession: A41502
A;Molecule type: protein
A;Residues: 1-13, 'S', 15-20 <HUA>
A;Cross-references: UNIPARC:UPI0000172DA4
A;Experimental source: strain UMR41
R;Evans Jr., D.J.; Evans, D.G.; Kirkpatrick, S.S.; Graham, D.Y.
Microb. Pathog. 10, 15-26, 1991
A;Title: Characterization of the Helicobacter pylori urease and purification of its subu
A;Reference number: A61371; MUID:91312104; PMID:1857197
A;Accession: A61371
A;Molecule type: protein
A;Residues: 1-5, 'R', 7-20, 'R', 22-30 <EVA>
A;Cross-references: UNIPARC:UPI0000172DA5
R;Turbett, G.R.; Hoj, P.B.; Horne, R.; Mee, B.J.
Infect. Immun. 60, 5259-5266, 1992
A;Title: Purification and characterization of the urease enzymes of Helicobacter species
A;Reference number: A49215; MUID:93084378; PMID:1452359
A;Accession: E49215
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-12 <TUR>
A;Cross-references: UNIPARC:UPI00000B1E50; PIDN:AB24359.1; PID:g260968
A;Experimental source: strain NCTC 11637
A;Note: sequence extracted from NCBI backbone (NCBIP:119486)
C;Genetics:
A;Gene: ureA; HP0073
C;Complex: heterodimer of 26K chain and 62K chain (see PIR:URKCBP)
C;Function:
A;Description: catalyzes the hydrolysis of one molecule of urea to carbon dioxide and tw
C;Superfamily: urease, fused gamma/beta subunit; urease 11K chain homology; urease 12K c
C;Keywords: heterodimer; hydrolase; metalloprotein
F;1-238/Product: urease 26K chain #status experimental <MAT>

F;1-101/Domain: urease 11K chain homology <U11>
F;108-204/Domain: urease 12K chain homology <U12>
Query Match 51.7%; Score 610.5; DB 1; Length 238;
Best Local Similarity 52.7%; Pred. No. 2.4e-43;
Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;
QY 1 VKLTPEQEKFLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLTPEKLDKMLHYAGELAKKKEKGIKLVYEAVALISAHIMEEARAGKKTAAELMQE 60
QY 61 CMHFLKDEVMGPNVDPDLGVATFPDGTGLVTVMWPIEPDEHFKAGEVKGFCDDKIE 120
Db 61 GRTLLKPDVMDGVASMIHEVGIEAMPFDGTGLVTVHTPIEANGKLVPGEL-FLKNEDIT 119
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFEANKALKFDREKAYGKRLLDIPSGNTLRIGAG 180
Db 120 INEGKAVSVKXNVGDRPVQIGSHFFHFEVFNCLDFDREKTFGRLLDIASGTAVRFPFG 179
QY 181 QTRKVLIPLGSGSKVIGMGLVNNIADERHKHKKALDKAKSHGF 224
Db 180 EEKSVELIDIGGNRRIFGFNALVDRQADNESKKIALHRAKERGF 223
RESULT 3
S35290
urease (EC 3.5.1.5) 26K chain - Helicobacter felis
N;Alternate names: urease alpha chain
C;Species: Helicobacter felis
C;Date: 31-Dec-1993 #sequence_revision 02-Dec-1994 #text_change 09-Jul-2004
C;Accession: S35290
R;Ferreiro, R.L.; Labigne, A.
Mol. Microbiol. 9, 323-333, 1993
A;Title: Cloning, expression and sequencing of Helicobacter felis urease genes.
A;Reference number: S35290; MUID:94018627; PMID:8412683
A;Accession: S35290
A;Molecule type: DNA
A;Residues: 1-237 <FER>
A;Cross-references: UNIPROT:Q08715; UNIPARC:UPI0000137D6C; EMBL:X69080; NID:g396160; PID
C;Genetics:
A;Gene: ureA
C;Complex: heterodimer
C;Function:
A;Description: catalyzes the hydrolysis of one molecule of urea to carbon dioxide and tw
C;Superfamily: urease, fused gamma/beta subunit; urease 11K chain homology; urease 12K c
C;Keywords: heterodimer; hydrolase; metalloprotein
F;1-101/Domain: urease 11K chain homology <U11>
F;108-204/Domain: urease 12K chain homology <U12>
Query Match 48.0%; Score 566.5; DB 1; Length 237;
Best Local Similarity 50.0%; Pred. No. 1.1e-39;
Matches 112; Conservative 43; Mismatches 68; Indels 1; Gaps 1;
QY 1 VKLTPEQEKFLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLTPEKLDKMLHYAGELAKKKEKGIKLVYEAVALISAHIMEEARAGKKTAAELMQE 60
QY 61 CMHFLKDEVMGPNVDPDLGVATFPDGTGLVTVMWPIEPDEHFKAGEVKGFCDDKIE 120
Db 61 GRTWLKKNVMDGVASMIHEVGIEANFPDGTGLVTIHTPVEDNGKLAPGEY-FLKNEDIT 119
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFEANKALKFDREKAYGKRLLDIPSGNTLRIGAG 180
Db 120 INAGKEAISLVKXNVGDRPVQVQSHFFHFEVFNCLDFDRAKSFCKRLDIASGTAVRFPFG 179
QY 181 QTRKVLIPLGSGSKVIGMGLVNNIADERHKHKKALDKAKSHGF 224
Db 180 EEKSVELIDIGGNRRIFGFNSLVDRQADADGKGLGLKRAKERGF 223
RESULT 4
D75586
urease, beta/gamma subunit - Deinococcus radiodurans (strain R1)

```

A;Reference number: JT0001
A;Accession: JY0001
A:Molecule type: protein
A;Residues: 1-246,'R','248-257','P','259-268','S','270-840 <TA2>'
A;Cross-references: UNIPARC:UPI0000172D9F
E;Mamiya, G.; Takishima, K.; Masakuni, M.; Kayumi, T.; Ogawa, K.
J. Protein Chem. 6, 55-59, 1987
A>Title: Complete amino acid sequence of jack bean urease.
A;Reference number: A60883
A;Accession: A60883
A:Molecule type: protein
A;Residues: 1-246,'R','248-840 <TA3>'
A;Cross-references: UNIPARC:UPI0000172D9E
E;Takishima, K.; Mamiya, G.
Protein Seq. Data Anal. 1, 103-106, 1987
A>Title: Location of the essential cysteine residue of jack bean urease.
A;Reference number: S03364; PMID:88190054; PMID:3447159
A;Accession: S03364
A:Molecule type: protein
A;Residues: 591-637 <TA4>'
A;Cross-references: UNIPARC:UPI0000172DA0
C;Comment: Each chain of the hexamer binds two nickel ions.
C;Superfamily: urease; urease 1LK chain homology; urease 12K chain homology; urease 12K chain homology; urease 62K chain homology; urease 62K chain homology; urease 11K chain homology; metallopeptide hydrolase; metalloprotein; nickel C;Keywords: homohexamer; hydrolase; metallopeptide; metallopeptide; Uil1>
F;1-101/Domain: urease 1LK chain homology <Uil1>
F;134-230/Domain: urease 12K chain homology <U12>
F;274-823/Domain: urease 62K chain homology <U62>
F;407,409,490,633/Binding site: nickel 2 (His, His) #status predicted
F;490,519,545/Binding site: nickel 1 (Lys, His, His) #status predicted
F;490,519,545/Binding site: carbon dioxide [lys] (covalent) (by urease activase) #status predicted
Query Match 37.2%; Score 439.5; DB 1; Length 840;
Best Local Similarity 39.6%; Pred. No. 1.7e+28;
Matches 99; Conservative 47; Mismatches 77; Indels 27; Gaps 3;

Qy 1 VKLPFKDEKEFLYYAGEVARRKRKAAGKLKNQPEAIAYISAHINDEARRGKTVAQLMEEE 60
   :|||:||| |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 1 MKLSPREVEKGLGNHAGLYLAQRARGRLNRYTEAVALLASQIMNEYARDGEKTVQAQMCL 60

Qy 61 CMHFLLKKDPGVGMVMVDLGAEATFPDDGTCLKVTNNPIEPDBEHFKAGEVKFG----- 114
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 61 GOHLGRQKVLPAPPHLLNAVQEAETFPDDGTCLKVTNHDPIS--RENGELOEQALFGSLPPVP 119

Qy 115 -----C-DKDIELNAGKEYTELVNTNEGPKSILHVSHGFHPFEANKA 154
   |::::::::::|::::::::::|::::::::::|::::::::::|::::::::::|
Db 120 SLDKFAETKEDNRIFGEIICEDECLTLNIGRKAKVLKVTSGDRPIQVGSHVFIVNPY 179

Qy 155 LKPDIKRAYGRKLDTSPGNLTIRIGACQTRKQVLIPLGGSKKIWMNLVNIAIDRRHKH 214
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 180 LTDFRRKAYGMRLNIATAAGTAVRFPGDCSVTLVSIEGNKVIKGNAIDGPNVTNLEA 239

Qy 215 ALDKRAKHGF 224
   |:|:|:|
Db 240 AMHAVRSKGF 249

RESULT 6
A96699
probable urease F12B7.10 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 05-Oct-2004
C;Accession: A96699
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L. Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A96141; PMID:21016719; PMID:11130712
```

[illegible]

submitted to the Brookhaven Protein Data Bank, June 1995

A:Reference number: A65977; PDB:1KRA
A:Contents: annotation; X-ray crystallography, 2.3 angstroms, residues 1-100
R:Jabri, E.; Carr, M.B.; Hausinger, R.P.; Karplus, P.A.
Science 268, 998-1004, 1995
A:Title: The crystal structure of urease from *Klebsiella aerogenes*.
A:Reference number: A56340; MUID:95273988; PMID:7754395
A:Contents: annotation; X-ray crystallography, 2.3 angstroms
C:Genetics:
A:Gene: ureA
C:Superfamily: urease, gamma subunit; urease 11K (gamma), 12K (beta), and
C:Complex: trimer of heterotrimeric subunits consisting of 11K (gamma), 12K (beta), and
C:Function:
A:Description: catalyzes the hydrolysis of one molecule of urea to carbon dioxide and two
C:Superfamily: urease, gamma subunit; urease 11K chain homology
C:Keywords: heterotrimer; hydrolase; metalloprotein
F:1-100/Domain: urease 11K chain homology <U1>

Query Match 23.8%; Score 281.5; DB 1; Length 100;
Best Local Similarity 59.0%; Pred. No. 1.7e-16;
Matches 59; Conservative 18; Mismatches 22; Indels 1; Gaps 1;

Qy 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIYISAHIMDEARRGKKTVAQLMEE 60
Db 1 MELTPREKDKLLFTAAALVAERRLARGLKLNYPESVALISAFIMEGARDG-KSVASLMEE 59

Qy 61 CMHFLKKDEVMGPNVGVNMPDLGVEATFPDGTKLVTNNWPI 100
Db 60 GRHVLREQVMGPEMIPDIQVEATFPDGSKLVTNNPI 99

RESULT 10

H70664

probable ureA protein - Mycobacterium tuberculosis (strain H37Rv)

C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:96334230
A:Accession: H70664
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-100 <COL>
C:Cross-references: UNIPROT:P50043; UNIPARC:UPI0000137D9C; GB:Z83859; GB:AL123456; NID:9
A:Experimental source: strain H37Rv
C:Genetics:
A:Gene: ureA
C:Superfamily: urease, gamma subunit; urease 11K chain homology
F:1-100/Domain: urease 11K chain homology <U1>

Query Match 23.6%; Score 278.5; DB 2; Length 100;
Best Local Similarity 58.0%; Pred. No. 3e-16;
Matches 58; Conservative 18; Mismatches 23; Indels 1; Gaps 1;

Qy 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIYISAHIMDEARRGKKTVAQLMEE 60
Db 1 MELTPHEQRLLLSYAAELARRRRARGRLNHPEATAVTADHILEGARDG-RTVAELMAS 59

Qy 61 CMHFLKKDEVMGPNVGVNMPDLGVEATFPDGTKLVTNNWPI 100
Db 60 GREVLGRDDVMGPEMLAEVQVEATFPDGTKLVTNNWPI 99

RESULT 11

DB5603

probable urease structural subunit A (gamma) [imported] - *Escherichia coli* (strain O157:
C:Species: *Escherichia coli*
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: DB5603; E85654

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: DB5603
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-100 <STO>
A:Cross-references: UNIPROT:Q8XAG2; UNIPARC:UPI00000D5DF7; GB:AE005174; NID:g12513945; P
A:Experimental source: strain O157:H7, substrain EDL933
A:Accession: E85654
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-100 <ST2>
A:Cross-references: UNIPARC:UPI00000D5DF7; GB:AE005174; NID:g12514462; PIDN:AAG55697.1;
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: ureA; ureA_2
C:Superfamily: urease, gamma subunit; urease 11K chain homology

Query Match 23.5%; Score 277.5; DB 2; Length 100;
Best Local Similarity 57.0%; Pred. No. 3.6e-16;
Matches 57; Conservative 20; Mismatches 22; Indels 1; Gaps 1;

Qy 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIYISAHIMDEARRGKKTVAQLMEE 60
Db 1 MELTPREKDKLLFTAAALVAERRLARGLKLNYPESVALISAFIMEGARDG-KSVAALMEE 59

Qy 61 CMHFLKKDEVMGPNVGVNMPDLGVEATFPDGTKLVTNNWPI 100
Db 60 GRHVLREQVMGPEMIPDIQVEATFPDGSKLVTNNPI 99

RESULT 12

B90794

urease gamma subunit [imported] - *Escherichia coli* (strain O157:H7, substrain RMD 050995
C:Species: *Escherichia coli*
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: B90794
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and geno
A:Reference number: A96299; MUID:21156231; PMID:11258796
A:Accession: B90794
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-100 <HAY>
A:Cross-references: UNIPROT:Q8XAG2; UNIPARC:UPI00000D5DF7; GB:BA000007; PIDN:BA834745.1;
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: ECol322
C:Superfamily: urease, gamma subunit; urease 11K chain homology

Query Match 23.5%; Score 277.5; DB 2; Length 100;
Best Local Similarity 57.0%; Pred. No. 3.6e-16;
Matches 57; Conservative 20; Mismatches 22; Indels 1; Gaps 1;

Qy 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIYISAHIMDEARRGKKTVAQLMEE 60
Db 1 MELTPREKDKLLFTAAALVAERRLARGLKLNYPESVALISAFIMEGARDG-KSVAALMEE 59

Qy 61 CMHFLKKDEVMGPNVGVNMPDLGVEATFPDGTKLVTNNWPI 100
Db 60 GRHVLREQVMGPEMIPDIQVEATFPDGSKLVTNNPI 99

RESULT 13

S74889

urease (EC 3.5.1.5) 11K chain - *Synechocystis* sp. (strain PCC 6803)
N:Alternate names: protein alr1256; urease gamma chain
C:Species: *Synechocystis* sp.

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 28, 2005, 07:57:42 ; Search time 86.8136 Seconds
(without alignments)
1836.687 Million cell updates/sec

Title: US-09-904-994B-2
Perfect score: 1181
Sequence: 1 VKLTPKEQKFLLYVAGEVA.....ADERHKKALDKAKSHGFIK 226

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------------|---------------------|
| 1 | 652 | 55.2 | 234 | 1 UREA_HELBI | Q8gh98 helicobacte |
| 2 | 622 | 52.7 | 234 | 1 UREA_HELHE | P42822 helicobacte |
| 3 | 613.5 | 51.9 | 238 | 1 UREA_HELPU | Q9zm24 helicobacte |
| 4 | 611.5 | 51.8 | 238 | 2 Q8RN05_HELPU | Q8rn05 helicobacte |
| 5 | 611.5 | 51.8 | 238 | 2 Q93NJ1_HELPU | Q93nj1 helicobacte |
| 6 | 610.5 | 51.7 | 238 | 1 UREA_HELPU | P14916 helicobacte |
| 7 | 610.5 | 51.7 | 238 | 2 Q93N13_HELPU | Q93ni3 helicobacte |
| 8 | 610.5 | 51.7 | 238 | 2 Q93N15_HELPU | Q93ni5 helicobacte |
| 9 | 609.5 | 51.6 | 238 | 2 Q93N18_HELPU | Q93ni8 helicobacte |
| 10 | 608.5 | 51.5 | 238 | 2 Q93M9_HELPU | Q93m9 helicobacte |
| 11 | 607.5 | 51.4 | 238 | 2 Q93285_HELPU | Q93285 helicobacte |
| 12 | 607.5 | 51.4 | 238 | 2 Q93N14_HELPU | Q93ni4 helicobacte |
| 13 | 607.5 | 51.4 | 238 | 2 Q93N10_HELPU | Q93ni0 helicobacte |
| 14 | 607.5 | 51.4 | 238 | 2 Q93N12_HELPU | Q93ni2 helicobacte |
| 15 | 605.5 | 51.3 | 238 | 2 Q84F76_HELPU | Q84f76 helicobacte |
| 16 | 604.5 | 51.2 | 238 | 2 Q93N19_HELPU | Q93ni9 helicobacte |
| 17 | 603.5 | 51.1 | 238 | 2 Q93N17_HELPU | Q93ni7 helicobacte |
| 18 | 603.5 | 51.1 | 238 | 2 Q933P5_HELPU | Q933p5 helicobacte |
| 19 | 602.5 | 51.0 | 238 | 2 Q93N10_HELPU | Q93ni0 helicobacte |
| 20 | 601.5 | 50.9 | 238 | 2 Q93N11_HELPU | Q93ni1 helicobacte |
| 21 | 601.5 | 50.9 | 238 | 2 Q93N18_HELPU | Q93ni8 helicobacte |
| 22 | 599.5 | 50.8 | 238 | 2 Q93N16_HELPU | Q93ni6 helicobacte |
| 23 | 599.5 | 50.8 | 238 | 2 Q93N17_HELPU | Q93ni7 helicobacte |
| 24 | 598.5 | 50.7 | 238 | 2 Q93N19_HELPU | Q93ni9 helicobacte |
| 25 | 596.5 | 50.5 | 238 | 1 UREA_HELFE | Q08715 helicobacte |
| 26 | 566.5 | 48.0 | 237 | 1 UREA_HELPU | P50044 helicobacte |
| 27 | 544.5 | 46.1 | 213 | 1 UREA_HELPU | Q7ped9 anophelies g |
| 28 | 542.5 | 45.9 | 347 | 2 Q7PED9_ANOGA | Q7ped9 helicobacte |
| 29 | 540.5 | 45.8 | 225 | 1 UREA_HELPU | Q972v9 sulfolobus |
| 30 | 498 | 42.2 | 219 | 1 UREA_HELPU | Q5f524 campylobact |
| 31 | 491 | 41.6 | 223 | 2 Q5FB24_CAMLA | |

| | | | | | |
|----|-------|------|-----|----------------|--------------------|
| 32 | 485 | 41.1 | 231 | 1 URE23_PRESM | Q883f3 pseudomonas |
| 33 | 483 | 40.9 | 228 | 1 URE23_DEIRA | Q9ryj3 deinococcus |
| 34 | 479 | 40.6 | 231 | 2 Q4ZUD1_PSESY | Q4zud1 pseudomonas |
| 35 | 444.5 | 37.6 | 838 | 2 Q7XAC5_SOYBN | Q7xac5 glycine max |
| 36 | 439.5 | 37.2 | 840 | 1 UREA_CANEN | P07374 canavalia e |
| 37 | 429.5 | 36.4 | 185 | 2 Q9R8A4_HELHP | Q9r8r4 helicobacte |
| 38 | 429 | 36.3 | 176 | 2 Q8KIY9_HELHE | Q8kiy9 helicobacte |
| 39 | 428 | 36.2 | 177 | 2 Q8KJ05_HELHE | Q8kj05 helicobacte |
| 40 | 427.5 | 36.2 | 185 | 2 Q9R8R2_HELHP | Q9r8r2 helicobacte |
| 41 | 426 | 36.1 | 164 | 2 Q8KI29_HELBI | Q8ki29 helicobacte |
| 42 | 426 | 36.1 | 172 | 2 Q8KHJ0_HELHE | Q8khj0 helicobacte |
| 43 | 426 | 36.1 | 179 | 2 Q8KI21_HELHE | Q8kiz1 helicobacte |
| 44 | 426 | 36.1 | 843 | 2 Q8W3L6_ORISA | Q8w3l6 oryza sativ |
| 45 | 425 | 36.0 | 190 | 2 Q6I6I6_CAMLA | Q6i6i6 campylobact |

ALIGNMENTS

RESULT 1

| | | | |
|------------|--|------|---------|
| UREA_HELBI | STANDARD; | PRT; | 234 AA. |
| ID | UREA_HELBI | | |
| AC | Q8GH98; | | |
| DT | 29-MAR-2004 (Rel. 43, Created) | | |
| DT | 29-MAR-2004 (Rel. 43, Last sequence update) | | |
| DT | 13-SEP-2005 (Rel. 48, Last annotation update) | | |
| DE | Urease alpha subunit (EC 3.5.1.5) (Urea amidohydrolase alpha subunit). | | |
| GN | Name=urea; | | |
| OS | Helicobacter bizzozeronii. | | |
| OC | Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; | | |
| OC | Helicobacteraceae; Helicobacter. | | |
| OX | NCBI_TaxID=56877; | | |
| RN | [1] | | |
| RP | NUCLEOTIDE SEQUENCE [GENOMIC DNA]. | | |
| RA | Zhu J., Chang Y.F.; | | |
| RT | "Cloning and characterization of Helicobacter bizzozeronii urease gene cluster."; | | |
| RL | Submitted (DEC-2000) to the EMBL/GenBank/DBAJ databases. | | |
| CC | -1- CATALYTIC ACTIVITY: Urea + H(2)O = CO(2) + 2 NH(3). | | |
| CC | -1- SUBUNIT: Heterodimer of an alpha subunit and a beta subunit (By similarity). | | |
| CC | -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity). | | |
| CC | -1- SIMILARITY: In the N-terminal section; belongs to the urease gamma subunit family. | | |
| CC | -1- SIMILARITY: In the C-terminal section; belongs to the urease beta subunit family. | | |
| CC | -1- CAUTION: In Helicobacter the alpha subunit is what is known, in other bacteria as the beta subunit. | | |
| CC | ----- | | |
| CC | This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed. | | |
| CC | ----- | | |
| DR | EMBL; AF330621; AA015373.1; -; Genomic_DNA. | | |
| DR | HSSP; P14916; 1592. | | |
| DR | SMR; Q8GH98; 1-224. | | |
| DR | HAMAP; MF_00739; fused; 1. | | |
| DR | InterPro; IPR002019; Urease_beta. | | |
| DR | InterPro; IPR002026; Urease_gamma_reg. | | |
| DR | InterPro; IPR008223; Urease_gamma_beta. | | |
| DR | Pfam; PF00699; Urease_beta; 1. | | |
| DR | Pfam; PF00547; Urease_gamma; 1. | | |
| DR | PIRSF; PIRSF001225; Urease_gamma_beta; 1. | | |
| DR | ProDom; PD002326; Urease_gamma; 1. | | |
| DR | ProDom; PD002319; Urease_gamma; 1. | | |
| DR | TIGRFAMS; TIGR00192; urease_beta; 1. | | |
| DR | TIGRFAMS; TIGR00193; urease_gamma; 1. | | |
| KW | Hydrolase. | | |
| FT | REGION 103 102 Urease_gamma. | | |
| FT | REGION 103 234 Urease_beta. | | |
| FT | SEQUENCE 234 AA; 25720 MW; 4567707785648843 CRC64; | | |
| FT | SEQUENCE 234 AA; 25720 MW; 4567707785648843 CRC64; | | |


```
DR InterPro; IPR002026; Urease_gamma_reg.
DR InterPro; IPR008223; Urease_gamma_beta.
DR Pfam; PF00699; Urease_beta; 1.
DR Pfam; PF00547; Urease_gamma; 1.
DR PIRSF; PIRSF001225; Urease_gamma_beta; 1.
DR ProDom; PD002326; Urease_beta; 1.
DR ProDom; PD002319; Urease_gamma; 1.
DR TIGRFAMs; TIGR00192; urease_beta; 1.
DR TIGRFAMs; TIGR00193; urease_gamma; 1.
KW Complete proteome; Hydrolase.
FT REGION 1 102 Urease gamma.
FT REGION 103 238 Urease beta.
SQ SEQUENCE 238 AA; 26568 MW; A10B9DC4156C0561 CRC64;

Query Match 51.9%; Score 613.5; DB 1; Length 238;
Best Local Similarity 53.1%; Pred. No. 1.7e-43;
Matches 119; Conservative 42; Mismatches 62; Indels 1; Gaps 1;

Qy 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLTPKELDKMLHYAGELAKRKEKGKILNYVEAVALISAHIMEEARAGKKTAAELMQE 60

Qy 61 CMHFLKDEVMGPNVDPGLGEATPPDGTGLVTNNWPIEPDEHFKAGEVKFGCDKDI 120
Db 61 GRTLLKPDVMDGVASMIHEVGIEAMPDPDGTGLVTHTPIEANGKLVPGEL-FLKNEDIT 119

Qy 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFEANKALFKDREKAYGKRLDIPSGNLTIRIGAG 180
Db 120 INEGKAVSVKVNVDGPRVQIGSHFFHFEVNRCLDFDREKTFGKRLDIASGTAVRFPFG 179

Qy 181 QTRKVLQIPLGSKKVIQMGNLVNNIADERHKKHAKLDKAKSHGF 224
Db 180 EEKSVELIDIGGNRRIFGFNALVDRQADNESKKTALHRAKERGF 223
```

```
RESULT 4
Q8RNU5 HELPY
ID Q8RNU5_HELPHY PRELIMINARY; PRT; 238 AA.
AC Q8RNU5;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Urease A subunit.
GN Name=ureA;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CH-CTX1;
RX MEDLINE=20123532; PubMed=10660136;
RA Opazo P., Muller I., Rollan A., Valenzuela P., Yudelevich A.,
RA Garcia-de la Guarda R., Urria S., Venegas A.;
RT "Serological response to Helicobacter pylori recombinant antigens in
RT Chilean infected patients with duodenal ulcer, non-ulcer dyspepsia and
RT gastric cancer.";
RL APMS 107:1069-1078 (1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CH-CTX1.
RX MEDLINE=22121498; PubMed=12125208;
RA Muller I., Medina-Selby A., Palacios J.L., Martinez P., Opazo P.,
RA Bruce E., Mancilla M., Valenzuela P., Yudelevich A., Venegas A.;
RT "Cloning and comparison of ten gene sequences of a Chilean H. pylori
RT strain with other H. pylori strains revealed higher variability for
RT VacA and CagA virulence factors.";
RL Biol. Res. 35:67-84 (2002).
DR EMBL; AF479027; AAL86897.1; -; Genomic_DNA.
DR HSSP; P14916; 1E9Y.
DR SNR; Q8RNU5; 1-238.
DR GO; GO:0005737; Cytoplasm; IEA.
DR GO; GO:0016151; F:nickel ion binding; IEA.
```

```
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0006807; P:nitrogen compound metabolism; IEA.
DR InterPro; IPR002019; Urease_beta.
DR InterPro; IPR008223; Urease_gamma_beta.
DR InterPro; IPR002026; Urease_gamma_reg.
DR Pfam; PF00699; Urease_beta; 1.
DR Pfam; PF00547; Urease_gamma; 1.
DR PIRSF; PIRSF001225; Urease_gamma_beta; 1.
DR ProDom; PD002326; Urease_beta; 1.
DR ProDom; PD002319; Urease_gamma; 1.
DR TIGRFAMs; TIGR00192; urease_beta; 1.
DR TIGRFAMs; TIGR00193; urease_gamma; 1.
SQ SEQUENCE 238 AA; 26710 MW; E934A6512A559B28 CRC64;

Query Match 51.8%; Score 611.5; DB 2; Length 238;
Best Local Similarity 52.7%; Pred. No. 2.5e-43;
Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;

Qy 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLTPKELDKMLHYAGELAKRKEKGKILNYVEAVRLISAHIMEEARGKKTAAELMQE 60

Qy 61 CMHFLKDEVMGPNVDPGLGEATPPDGTGLVTNNWPIEPDEHFKAGEVKFGCDKDI 120
Db 61 GRTLLKPDVMDGVASMIHEVGIEAMPDPDGTGLVTHTPIEANGKLVPGEL-FLKNEDIT 119

Qy 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFEANKALFKDREKAYGKRLDIPSGNLTIRIGAG 180
Db 120 INEGKAVSVKVNVDGPRVQIGSHFFHFEVNRCLDFDREKTFGKRLDIASGTAVRFPFG 179

Qy 181 QTRKVLQIPLGSKKVIQMGNLVNNIADERHKKHAKLDKAKSHGF 224
Db 180 EEKSVELIDIGGNRRIFGFNALVDRQADNESKKTALHRAKERGF 223

RESULT 5
Q93NJ1 HELPY
ID Q93NJ1_HELPHY PRELIMINARY; PRT; 238 AA.
AC Q93NJ1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Urease protein Urea.
GN Name=ureA;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Chu C., Ou J.T.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF373558; AAK69724.1; -; Genomic_DNA.
DR HSSP; P14916; 1E9Y.
DR SNR; Q93NJ1; 1-238.
DR GO; GO:0005737; Cytoplasm; IEA.
DR GO; GO:0016151; F:nickel ion binding; IEA.
DR GO; GO:0006807; P:urease activity; IEA.
DR GO; GO:0009039; F:nitrogen compound metabolism; IEA.
DR InterPro; IPR002019; Urease_beta.
DR InterPro; IPR008223; Urease_gamma_reg.
DR InterPro; IPR002026; Urease_gamma_beta; 1.
DR Pfam; PF00699; Urease_beta; 1.
DR Pfam; PF00547; Urease_gamma; 1.
DR PIRSF; PIRSF001225; Urease_gamma_beta; 1.
DR ProDom; PD002326; Urease_beta; 1.
DR ProDom; PD002319; Urease_gamma; 1.
DR TIGRFAMs; TIGR00192; urease_beta; 1.
DR TIGRFAMs; TIGR00193; urease_gamma; 1.
SQ SEQUENCE 238 AA; 26672 MW; EE82BF861D03BA50 CRC64;

Query Match 51.8%; Score 611.5; DB 2; Length 238;
Best Local Similarity 53.1%; Pred. No. 2.5e-43;
```


FT TURN 67 69
FT STRAND 70 70
FT TURN 72 73
FT HELIX 74 77
FT STRAND 80 87
FT TURN 88 89
FT STRAND 90 97
FT TURN 108 109
FT STRAND 111 111
FT STRAND 117 119
FT TURN 120 121
FT STRAND 128 133
FT STRAND 139 142
FT TURN 143 144
FT HELIX 147 149
FT TURN 152 153
FT STRAND 154 155
FT HELIX 158 161
FT TURN 162 163
FT STRAND 164 166
FT TURN 170 171
FT STRAND 173 176
FT TURN 178 179
FT STRAND 181 188
FT TURN 191 192
FT STRAND 194 195
FT TURN 198 199
FT STRAND 204 205
FT HELIX 208 221
FT TURN 222 222
FT TURN 224 225
SQ SEQUENCE 238 AA; 4E77328669CD9A2D CRC64;

Query Match 51.7%; Score 610.5; DB 1; Length 238;
Best Local Similarity 52.7%; Pred. No. 3e-43;
Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;

Qy 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEATAYISAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLTPKELDKMLHYAGELAKRKKEGKIKLYVEAVALISAHIMEEARAGKKSAAELMQE 60

Qy 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTCLVTNNWPIEDPDEHFKAGEVKGCDKDIE 120
Db 61 GRTLLKPDVMDGVASMIHEVGIEAMFPDGTCLVTVHTPIEANGKLVPGEL-FLKNEDIT 119

Qy 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFEANKALKPDREKAYGKRLDIPSGNTLRIGAG 180
Db 120 INEGKXAVSVKXNVGDRPVQIGSHFFHFEVNRCLDFDREKTFGKRLDIASGTAVRPEPG 179

Qy 181 QTRKVLQIPLGGSKKVGIMGLVNNIADERHKHAKLDKAKSHGF 224
Db 180 EKSVELIDIGNRRIFGFNALVDRQADNESKXIALHRAKERGF 223

RESULT 7
Q93NI3_HELPY
ID Q93NI3_HELPY PRELIMINARY; PRT; 238 AA.
AC Q93NI3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Urease protein Urea.
GN Name=ureA;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Chu C., Ou J.T.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF373570; AAK69736.1; -, Genomic_DNA.
DR HSSP; P14916; IE9Y.

DR SMR; Q93NI3; 1-238.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0016151; F:nickel ion binding; IEA.
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0006807; P:nitrogen compound metabolism; IEA.
DR InterPro; IPR002019; Urease_beta.
DR InterPro; IPR008223; Urease_gamma.
DR InterPro; IPR002026; Urease_gamma_reg.
DR Pfam; PF00699; Urease_beta; 1.
DR Pfam; PF00547; Urease_gamma; 1.
DR PIRSF; PIRSF001225; Urease_gamma_beta; 1.
DR ProDom; PD002326; Urease_beta; 1.
DR ProDom; PD002319; Urease_gamma; 1.
DR TIGRFAMs; TIGR00192; urease_beta; 1.
DR TIGRFAMs; TIGR00193; urease_gamma; 1.
SQ SEQUENCE 238 AA; 26586 MW; DFDBA17D016FC351 CRC64;

Query Match 51.7%; Score 610.5; DB 2; Length 238;
Best Local Similarity 53.1%; Pred. No. 3e-43;
Matches 119; Conservative 42; Mismatches 62; Indels 1; Gaps 1;

Qy 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAVISAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLTPKELDKMLHYAGELAKRKKEGKIKLYVEAVALISAHIMEEARAGKKSAAELMQE 60

Qy 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTCLVTNNWPIEDPDEHFKAGEVKGCDKDIE 120
Db 61 GRTLLKPDVMDGVASMIHEVGIEAMFPDGTCLVTVHTPIEANGKLVPGEL-FLKNEDIT 119

Qy 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFEANKALKPDREKAYGKRLDIPSGNTLRIGAG 180
Db 120 INEGKXAVSVKXNVGDRPVQIGSHFFHFEVNRCLDFDREKTFGKRLDIASGTAVRPEPG 179

Qy 181 QTRKVLQIPLGGSKKVGIMGLVNNIADERHKHAKLDKAKSHGF 224
Db 180 EKSVELIDIGNRRIFGFNALVDRQADNESKXIALHRAKERGF 223

RESULT 8
Q93NI5_HELPY
ID Q93NI5_HELPY PRELIMINARY; PRT; 238 AA.
AC Q93NI5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Urease protein Urea.
GN Name=ureA;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Chu C., Ou J.T.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF373568; AAK69734.1; -, Genomic_DNA.
DR HSSP; P14916; IE9Y.
DR SMR; Q93NI5; 1-238.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0016151; F:nickel ion binding; IEA.
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0006807; P:nitrogen compound metabolism; IEA.
DR InterPro; IPR002019; Urease_beta.
DR InterPro; IPR008223; Urease_gamma.
DR InterPro; IPR002026; Urease_gamma_reg.
DR Pfam; PF00699; Urease_beta; 1.
DR Pfam; PF00547; Urease_gamma; 1.
DR PIRSF; PIRSF001225; Urease_gamma_beta; 1.
DR ProDom; PD002326; Urease_beta; 1.
DR ProDom; PD002319; Urease_gamma; 1.
DR TIGRFAMs; TIGR00192; urease_beta; 1.
DR TIGRFAMs; TIGR00193; urease_gamma; 1.
SQ SEQUENCE 238 AA; 26539 MW; 251AD890042A4262 CRC64;


```

DR GO: 0005737; C:cytoplasm; IEA.
DR GO: 00016151; F:nickel ion binding; IEA.
DR GO: 0009039; F:nitrogen activity; IEA.
DR GO: 0006807; P:nitrogen compound metabolism; IEA.
DR InterPro: IPR002019; Urease_beta.
DR InterPro: IPR008223; Urease_gamma_beta.
DR InterPro: IPR002026; Urease_gamma_reg.
DR Pfam: PF00699; Urease_beta; 1.
DR Pfam: PF00547; Urease_gamma; 1.
DR PIRSF: PIRSF001225; Urease_gamma_beta; 1.
DR ProDom: PD002326; Urease_beta; 1.
DR ProDom: PD002319; Urease_gamma; 1.
DR TIGRFAMs: TIGR00192; urease_beta; 1.
DR TIGRFAMs: TIGR00193; urease_gamma; 1.
SQ SEQUENCE 238 AA; 26511 MW; CA6677D2788BDD2E CRC64;

Query Match 51.4%; Score 607.5; DB 2; Length 238;
Best Local Similarity 52.2%; Pred. No. 5.4e-43;
Matches 117; Conservative 44; Mismatches 62; Indels 1; Gaps 1

Qy 1 VKLTPKQEKFLLYYAGEVARKKAEGKLKNQPEAIAVISAHIMDEARGGKTVQALMEE 60
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1 MKLTPKELDKLMLHYAGELAKRKEGKILNYYVEAVALISAHIMEEARAGKKSAAELMQE 60

Qy 61 CMHFLKDEVPVGVNMPDLGVETFPDGTKLVTVMWPIEPDEHFKAGEVFKGCDKOIE 120
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
61 GRSLKPDVMDGVASMIHEVIGIEAMPFDGTKLVTVHTPIEANGKLVPGEI-FLKNEDIT 119

Qy 121 LNAQKEVTELEVNEGPKSLHVGSHPFFHEFANKALKPDREKAYGKRLDIPSGNTLIRIGAG 180
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
120 INEGKAVSVKXVNVGDRPVOIGSHFFHFEVNRCLDPDREKTFGKRLDIASGTAVRFEFG 179

Qy 181 QTRKVLQIPGGSKVIGMNLVNNIADERHKHKALDKAKSHGF 224
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
180 EKSVELIDIGNRRIFGFNALVDRQADNESKKTALHEAKERGF 223

RESULT 13
Q93NJ0_HELPY
ID Q93NJ0_HELPY PRELIMINARY; PRT; 238 AA.
AC Q93NJ0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Urease protein Urea.
GN Name=urea;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RX [1]
RP NUCLEOTIDE SEQUENCE.
RL Chu C., Ou J.T.;
RA Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF373560; AAK69726.1; -; Genomic_DNA.
DR HSSP; P14916; 1E9Y.
DR SMR; Q93NJ0; 1-238.
DR GO: 0005737; C:cytoplasm; IEA.
DR GO: 00016151; F:nickel ion binding; IEA.
DR GO: 0009039; F:nitrogen activity; IEA.
DR GO: 0006807; P:nitrogen compound metabolism; IEA.
DR InterPro: IPR002019; Urease_beta.
DR InterPro: IPR008223; Urease_gamma_beta.
DR InterPro: IPR002026; Urease_gamma_reg.
DR Pfam: PF00699; Urease_beta; 1.
DR Pfam: PF00547; Urease_gamma; 1.
DR PIRSF: PIRSF001225; Urease_gamma_beta; 1.
DR ProDom: PD002326; Urease_beta; 1.
DR ProDom: PD002319; Urease_gamma; 1.
DR TIGRFAMs: TIGR00192; urease_beta; 1.
DR TIGRFAMs: TIGR00193; urease_gamma; 1.
SQ SEQUENCE 238 AA; 26629 MW; 5C082660AAC88469 CRC64;

```

```

Qy      181 QTRKRVQLIPLGGSKKVTGMGLVNNIADERHHKALDKAKSHGF   224
       : : ||| :||::: | ||: | | | | | | | |
Db      180 EEKSVELIDIGGNRRIRFGFNALVDROADNESKTIALHRAKERGF   223

RESULT 15
Q84F76_HELPY
ID    Q84F76_HELPY PRELIMINARY;          PET;   238 AA.
AC    Q84F76;
DT    01-JUN-2003 (TrEMBLrel. 24, Created)
DD    01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT    01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE    Urease alpha (EC 3.5.1.5).
GN    Name=urea;
OS    Helicobacter pylori (Campylobacter pylori).
OC    Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC    Helicobacteraceae; Helicobacter.
OX    NCBI_TaxID=210;
RN    [1]

```

| | |
|----|---|
| RA | Kirami Arzenani M., Mohammadi M.; |
| RA | Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases. |
| RL | EMBL; AY272442; AA034402.1; -; Genomic DNA. |
| DR | HSSP; P14916; IESY. |
| DR | SMP; Q84F76; 1-238. |
| DR | GO; GO:0005737; Cytoplasm; IEA. |
| DR | GO; GO:0016787; F:hydrolase activity; IEA. |
| DR | GO; GO:0016151; F:nickel ion binding; IEA. |
| DR | GO; GO:0009039; F:urease activity; IEA. |
| DR | GO; GO:0006807; P:nitrogen compound metabolism; IEA. |
| DR | InterPro; IPRO02019; Urease_beta. |
| DR | InterPro; IPRO08223; Urease_gammabeta. |
| DR | InterPro; IPRO02026; Urease_gamma_reg. |
| DR | Pfam; PF00699; Urease_beta; 1. |
| DR | Pfam; PF00547; Urease_gamma; 1. |
| DR | PIRSP; PIRSP001225; Urease_gammabeta; 1. |
| DR | ProDom; PD002326; Urease_beta; 1. |
| DR | ProDom; PD002319; Urease_gamma; 1. |
| DR | TIGRFAMs; TIGR00192; urease_beta; 1. |
| DR | TIGRFAMs; TIGR00193; urease_gamma; 1. |
| KW | Hydrolase. |
| SQ | SEQUENCE 238 AA; 26491 MW; 6384770531FCB38C CRC64; |
| | Query Match 51.3%; Score 605.5; DB 2; Length 238; |
| | Best Local Similarity 52.7%; Pred No. 8e-43; |
| | Matches 118; Conservative 42; Mismatches 63; Indels 1; Gaps 1; |
| Qy | 1 VKLTPKQEKEFLLYAGSEVARRKKAEGKLKNOPAIYAISAHIMDEARRGKTVAQLMEE 60 |
| Db | 1 MKLTPEKLDKMLHYAGELAKKRKEGKIKNVEAVALLSAHIMEEARAGKKTAALMQE 60 |
| Qy | 61 CMHFLLKDEVGVGNVPVDLGVEATFPDGTCLVTNVWPIDPDHFHKAGEVKFGCDKDIE 120 |
| Db | 61 GRITLLKPPDVNDVGASMIHEVGIEAMFPDGTCLVTVTHTPIEANGKLVPGEI-LFKNEDIT 119 |

```

QY 121 LNAKGVIEUWVINEGFASLNVQSHFAPFEAKVAKLNDIAEGLGKLUUFGUNLAKIGAS 168
Db 120 INEGKKAWSVKVKNVGRPVQVGSHLHFFEVNRCLEDFDREKTFGKRLDIASGTAVRFBFG 179
QY 181 QTRKVQLIPLGGSKKVIGMNGLVNNIADERHKHKALDKAKSHGF 224
Db 180 EEKSVELIDIGNRRIFGFNALVDROADNESKKTALHRAKERGF 223

Search completed: November 28, 2005, 08:19:44
Job time : 88.8136 secs

```

Search completed: November 28, 2005, 08:19:44
Job time : 88.8136 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 28, 2005, 07:55:02 ; Search time 73.7204 Seconds
(without alignments)
1346.976 Million cell updates/sec

Title: US-09-904-994B-2

Perfect score: 1181

Sequence: 1 VKLTPKEQKFLLYAGEVA.....ADRRHKKALDKAKSHGFIK 226

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 43937871 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A_Geneseq_21.*
- 1: Geneseq1980s.*
 - 2: Geneseq1990s.*
 - 3: Geneseq2000s.*
 - 4: Geneseq2001s.*
 - 5: Geneseq2002s.*
 - 6: Geneseq2003as.*
 - 7: Geneseq2003bs.*
 - 8: Geneseq2004s.*
 - 9: Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 1181 | 100.0 | 226 | 5 | Adj58238 Urease su |
| 2 | 1178 | 99.7 | 226 | 5 | Adj58250 UreaseXY |
| 3 | 1174 | 99.4 | 226 | 5 | Adj58241 UreaseXY |
| 4 | 1174 | 99.4 | 226 | 5 | Adj58247 UreaseXY |
| 5 | 1170 | 99.1 | 226 | 5 | Adj58244 UreaseXY |
| 6 | 652 | 55.2 | 234 | 8 | Adq37849 H. bizoz |
| 7 | 613.5 | 51.9 | 238 | 4 | AARQ37849 H. bizoz |
| 8 | 610.5 | 51.7 | 238 | 2 | AARQ4579 Part of p |
| 9 | 610.5 | 51.7 | 238 | 2 | AARQ4579 Part of p |
| 10 | 610.5 | 51.7 | 238 | 4 | AARQ4579 Part of p |
| 11 | 610.5 | 51.7 | 238 | 5 | AARQ4579 Part of p |
| 12 | 610.5 | 51.7 | 245 | 4 | AEB55115 Urease A |
| 13 | 607.5 | 51.4 | 238 | 8 | ADS09178 H. pylori |
| 14 | 603.5 | 51.1 | 238 | 8 | ADM28643 Helicobac |
| 15 | 602.5 | 51.0 | 238 | 2 | AAR67375 H. pylori |
| 16 | 596.5 | 50.5 | 238 | 5 | AAM49645 Salmonell |
| 17 | 595.5 | 50.4 | 238 | 2 | AAR12515 A subunit |
| 18 | 573.5 | 48.6 | 228 | 3 | AAB52559 Helicobac |
| 19 | 573.5 | 48.6 | 228 | 3 | AAB52559 Helicobac |
| 20 | 572.5 | 48.5 | 237 | 2 | AAR74336 Helicobac |
| 21 | 572.5 | 48.5 | 237 | 2 | AAR74336 Helicobac |
| 22 | 572.5 | 48.5 | 806 | 2 | AAR67371 Urease A |
| 23 | 562.5 | 47.6 | 222 | 5 | AAM49647 Gerbil ur |
| 24 | 559.5 | 47.4 | 219 | 5 | AAM49646 Murine ur |

| | | | | | | |
|----|-------|------|------|---|-----------|--------------------|
| 25 | 559.5 | 47.4 | 224 | 5 | AAM49648 | Aam49648 Murine ur |
| 26 | 502.5 | 42.5 | 811 | 6 | ABU42650 | Abu42650 Protein e |
| 27 | 497.5 | 42.1 | 207 | 6 | ABU42523 | Abu42523 Protein e |
| 28 | 493 | 41.7 | 779 | 6 | ABU41210 | Abu41210 Protein e |
| 29 | 489.5 | 41.4 | 226 | 4 | AAU33952 | Aau33952 Staphyloc |
| 30 | 477 | 40.4 | 207 | 6 | ABU20285 | Abu20285 Protein e |
| 31 | 470.5 | 39.8 | 209 | 6 | ABU40327 | Abu40327 Protein e |
| 32 | 470 | 39.8 | 261 | 6 | ABU27583 | Abu27583 Protein e |
| 33 | 469.5 | 39.8 | 207 | 6 | ABU17114 | Abu17114 Protein e |
| 34 | 455.5 | 38.6 | 228 | 6 | ABU22153 | Abu22153 Protein e |
| 35 | 449.5 | 38.1 | 209 | 2 | AAR67376 | Aar67376 P. mirabi |
| 36 | 443.5 | 37.6 | 1216 | 2 | AAR67376 | Aar67376 P. mirabi |
| 37 | 439.5 | 37.2 | 840 | 8 | ADJ66032 | Adj66032 Coil form |
| 38 | 439.5 | 37.2 | 840 | 9 | AEA54210 | Aea54210 C. ensifo |
| 39 | 437 | 37.0 | 161 | 5 | AEAS0981 | Aeas0981 Helicobac |
| 40 | 429.5 | 36.4 | 270 | 2 | AAR67377 | Aar67377 Canavalia |
| 41 | 368.5 | 31.2 | 837 | 8 | ADG65522 | Adg65522 C. posada |
| 42 | 281.5 | 23.8 | 100 | 2 | AAW37775 | Aaw37775 Klebsiell |
| 43 | 281.5 | 23.8 | 100 | 3 | AAAY81823 | Aay81823 Klebsiell |
| 44 | 281.5 | 23.8 | 100 | 4 | AAU36152 | Aau36152 Klebsiell |
| 45 | 281.5 | 23.8 | 100 | 6 | ABU32326 | Abu32326 Protein e |

ALIGNMENTS

RESULT 1
ID ADJ58238 standard; protein; 226 AA.
XX
AC ADJ58238;
XX
DT 06-MAY-2004 (first entry)
XX
DE Urease subunit polypeptide X.
XX
KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.
XX
OS Helicobacter felis.
XX
FN EPI176192-A2.
PD 30-JAN-2002.
XX
PF 11-JUL-2001; 2001EP-00202666.
XX
PR 17-JUL-2000; 2000EP-00202565.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Kusters JG, Cattoli G;
XX
DR WPI; 2002-124384/17.
XX
DR N-PSDB; ADJ58237.
XX
PT Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
PT the diagnosis of Helicobacter felis infections and in the preparation of
PT vaccines.
XX
PS Claim 8; SEQ ID NO 2; 76pp; English.
XX
CC The present invention relates to a novel Helicobacter felis urease X and
CC Y subunit polypeptides and immunogenic fragments. The polypeptides are
CC used in the manufacture of vaccines against Helicobacter felis infections
CC and in diagnostic tests to detect antibodies against Helicobacter felis.
CC Helicobacter felis is difficult to grow so it is more convenient to use
CC the expression products of the genes encoding the urease X and Y subunits
CC in the manufacture of vaccines. The present sequence represents an urease
CC X subunit polypeptide of the invention.
XX
SQ Sequence 226 AA;
Query Match 100.0%; Score 1181; DB 5; Length 226;

```
Best Local Similarity 100.0%; Pred. No. 6.8e-116;
Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKLTPKEQKFLLYAGVARRKKAEGKLNOPEALAYISAHIMDEARRGKKTVAQLMEE 60
DB 1 VKLTPKEQKFLLYAGVARRKKAEGKLNOPEALAYISAHIMDEARRGKKTVAQLMEE 60
QY 61 CMHFLKKDEVMGPGVGNMVPDLGVEATFPDGTGTLVTNNWPIEPDEHFKAQYKRLDIPSGNTRLRIGAG 120
DB 61 CMHFLKKDEVMGPGVGNMVPDLGVEATFPDGTGTLVTNNWPIEPDEHFKAQYKRLDIPSGNTRLRIGAG 120
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDEKAYGKRLDIPSGNTRLRIGAG 180
DB 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDEKAYGKRLDIPSGNTRLRIGAG 180
QY 181 QTRKVQLIPLGGSKKVIWGNGLVNNIADERRHKHAKLDKAKSHGFIK 226
DB 181 QTRKVQLIPLGGSKKVIWGNGLVNNIADERRHKHAKLDKAKSHGFIK 226

RESULT 2
ADJ58250
ID ADJ58250 standard; protein; 226 AA.
XX
AC ADJ58250;
XX
DT 06-MAY-2004 (first entry)
XX
DE UreaseXY subunit #7.
XX
KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.
XX
OS Helicobacter felis.
XX
PN EP1176192-A2.
XX
PD 30-JAN-2002.
XX
PF 11-JUL-2001; 2001EP-00202666.
XX
PR 17-JUL-2000; 2000EP-00202565.
XX
PA (ALKU ) AKZO NOBEL NV.
XX
PI Kusters JG, Cattoli G;
XX
DR WPI; 2002-124384/17.
XX
DR N-PSDB; ADJ58249.
XX
PT Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
PT the diagnosis of Helicobacter felis infections and in the preparation of
PT vaccines.
XX
PS Disclosure; SEQ ID NO 14; 76pp; English.
XX
CC The present invention relates to a novel Helicobacter felis urease X and
CC Y subunit polypeptides and immunogenic fragments. The polypeptides are
CC used in the manufacture of vaccines against Helicobacter felis infections
CC and in diagnostic tests to detect antibodies against Helicobacter felis.
CC Helicobacter felis is difficult to grow so it is more convenient to use
CC the expression products of the genes encoding the urease X and Y subunits
CC in the manufacture of vaccines. The present sequence represents an
CC ureaseXY subunit of the invention.
XX
SQ Sequence 226 AA;
XX
Query Match 99.7%; Score 1178; DB 5; Length 226;
Best Local Similarity 99.6%; Pred. No. 1.4e-115;
Matches 225; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKLTPKEQKFLLYAGVARRKKAEGKLNOPEALAYISAHIMDEARRGKKTVAQLMEE 60
DB 1 VKLTPKEQKFLLYAGVARRKKAEGKLNOPEALAYISAHIMDEARRGKKTVAQLMEE 60
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDEKAYGKRLDIPSGNTRLRIGAG 180
DB 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDEKAYGKRLDIPSGNTRLRIGAG 180
QY 181 QTRKVQLIPLGGSKKVIWGNGLVNNIADERRHKHAKLDKAKSHGFIK 226
DB 181 QTRKVQLIPLGGSKKVIWGNGLVNNIADERRHKHAKLDKAKSHGFIK 226

RESULT 3
ADJ58241
ID ADJ58241 standard; protein; 226 AA.
XX
AC ADJ58241;
XX
DT 06-MAY-2004 (first entry)
XX
DE UreaseXY subunit #1.
XX
KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.
XX
OS Helicobacter felis.
XX
PN EP1176192-A2.
XX
PD 30-JAN-2002.
XX
PF 11-JUL-2001; 2001EP-00202666.
XX
PR 17-JUL-2000; 2000EP-00202565.
XX
PA (ALKU ) AKZO NOBEL NV.
XX
PI Kusters JG, Cattoli G;
XX
DR WPI; 2002-124384/17.
XX
DR N-PSDB; ADJ58240.
XX
PT Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
PT the diagnosis of Helicobacter felis infections and in the preparation of
PT vaccines.
XX
PS Disclosure; SEQ ID NO 5; 76pp; English.
XX
CC The present invention relates to a novel Helicobacter felis urease X and
CC Y subunit polypeptides and immunogenic fragments. The polypeptides are
CC used in the manufacture of vaccines against Helicobacter felis infections
CC and in diagnostic tests to detect antibodies against Helicobacter felis.
CC Helicobacter felis is difficult to grow so it is more convenient to use
CC the expression products of the genes encoding the urease X and Y subunits
CC in the manufacture of vaccines. The present sequence represents an
CC ureaseXY subunit of the invention.
XX
SQ Sequence 226 AA;
XX
Query Match 99.4%; Score 1174; DB 5; Length 226;
Best Local Similarity 99.1%; Pred. No. 3.7e-115;
Matches 224; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VKLTPKEQKFLLYAGVARRKKAEGKLNOPEALAYISAHIMDEARRGKKTVAQLMEE 60
DB 1 VKLTPKEQKFLLYAGVARRKKAEGKLNOPEALAYISAHIMDEARRGKKTVAQLMEE 60
QY 61 CMHFLKKDEVMGPGVGNMVPDLGVEATFPDGTGTLVTNNWPIEPDEHFKAQYKRLDIPSGNTRLRIGAG 120
DB 61 CMHFLKKDEVMGPGVGNMVPDLGVEATFPDGTGTLVTNNWPIEPDEHFKAQYKRLDIPSGNTRLRIGAG 120
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDEKAYGKRLDIPSGNTRLRIGAG 180
DB 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDEKAYGKRLDIPSGNTRLRIGAG 180
```

```
Db 121 LNAGKEVTELVNTEGPKSLHVGSHFFPETNKALKAFDREKAYGKRLDIPSGNTRLRIGAG 180
Qy 181 QTRKQVLIPLGSKKVIQGNGLVNNIADERHKHAKLDKAKSHGFIK 226
Db 181 QTRKQVLIPLGSKKVIQGNGLVNNIADERHKHAKLDKAKSHGFIK 226

RESULT 4
ADJ58247
ID ADJ58247 standard; protein; 226 AA.
XX
AC ADJ58247;
XX
DT 06-MAY-2004 (first entry)
DE UreaseXY subunit #5.
XX
KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.
XX
OS Helicobacter felis.
XX
PN EP1176192-A2.
XX
PD 30-JAN-2002.
XX
PF 11-JUL-2001; 2001EP-00202666.
XX
PR 17-JUL-2000; 2000EP-00202565.
XX
PA (ALKU ) AKZO NOBEL NV.
XX
PI Kusters JG, Cattoli G;
XX
PT Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
DR WPI; 2002-124384/17.
DR N-PSDB; ADJ58246.
XX
PS Disclosure; SEQ ID NO 11; 76pp; English.
XX
CC The present invention relates to a novel Helicobacter felis urease X and
CC Y subunit polypeptides and immunogenic fragments. The polypeptides are
CC used in the manufacture of vaccines against Helicobacter felis infections
CC and in diagnostic tests to detect antibodies against Helicobacter felis.
CC Helicobacter felis is difficult to grow so it is more convenient to use
CC the expression products of the genes encoding the urease X and Y subunits
CC in the manufacture of vaccines. The present sequence represents an
CC ureaseXY subunit of the invention.
XX
SQ Sequence 226 AA;

Query Match 99.4%; Score 1174; DB 5; Length 226;
Best Local Similarity 99.1%; Pred. No. 3.7e-115;
Matches 224; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
Db 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAELMEE 60
Qy 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTCLVTNNWPIEPDEHFKAGEVFGCDKDIE 120
Db 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTCLVTNNWPIEPDEHFKAGEVFGCDKDIE 120
Qy 121 LNAGKEVTELVNTEGPKSLHVGSHFFPETNKALKAFDREKAYGKRLDIPSGNTRLRIGAG 180
Db 121 LNAGKEVTELVNTEGPKSLHVGSHFFPETNKALKAFDREKAYGKRLDIPSGNTRLRIGAG 180
Qy 181 QTRKQVLIPLGSKKVIQGNGLVNNIADERHKHAKLDKAKSHGFIK 226
Db 181 QTRKQVLIPLGSKKVIQGNGLVNNIADERHKHAKLDKAKSHGFIK 226
```

```
RESULT 5
ADJ58244
ID ADJ58244 standard; protein; 226 AA.
XX
AC ADJ58244;
XX
DT 06-MAY-2004 (first entry)
DE UreaseXY subunit #3.
XX
KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.
XX
OS Helicobacter felis.
XX
PN EP1176192-A2.
XX
PD 30-JAN-2002.
XX
PF 11-JUL-2001; 2001EP-00202666.
XX
PR 17-JUL-2000; 2000EP-00202565.
XX
PA (ALKU ) AKZO NOBEL NV.
XX
PI Kusters JG, Cattoli G;
XX
PT Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
DR WPI; 2002-124384/17.
DR N-PSDB; ADJ58243.
XX
PS Disclosure; SEQ ID NO 8; 76pp; English.
XX
CC The present invention relates to a novel Helicobacter felis urease X and
CC Y subunit polypeptides and immunogenic fragments. The polypeptides are
CC used in the manufacture of vaccines against Helicobacter felis infections
CC and in diagnostic tests to detect antibodies against Helicobacter felis.
CC Helicobacter felis is difficult to grow so it is more convenient to use
CC the expression products of the genes encoding the urease X and Y subunits
CC in the manufacture of vaccines. The present sequence represents an
CC ureaseXY subunit of the invention.
XX
SQ Sequence 226 AA;

Query Match 99.1%; Score 1170; DB 5; Length 226;
Best Local Similarity 98.7%; Pred. No. 9.8e-115;
Matches 223; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
Db 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAELMEE 60
Qy 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTCLVTNNWPIEPDEHFKAGEVFGCDKDIE 120
Db 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTCLVTNNWPIEPDEHFKAGEVFGCDKDIE 120
Qy 121 LNAGKEVTELVNTEGPKSLHVGSHFFPETNKALKAFDREKAYGKRLDIPSGNTRLRIGAG 180
Db 121 LNAGKEVTELVNTEGPKSLHVGSHFFPETNKALKAFDREKAYGKRLDIPSGNTRLRIGAG 180
Qy 181 QTRKQVLIPLGSKKVIQGNGLVNNIADERHKHAKLDKAKSHGFIK 226
Db 181 QTRKQVLIPLGSKKVIQGNGLVNNIADERHKHAKLDKAKSHGFIK 226

RESULT 6
ADQ37849
ID ADQ37849 standard; protein; 234 AA.
XX
```


Db 1 MKLTPKELDKMLHYAGELARRKKEGKILNYYVEAVALISAHIMEEARAGKKTAAELMQE 60
QY 61 CMHFLKDEVMFGVGNMVPDLGVEATFPDGTGLVTNNWPIEPDEHFKAGVEKFGCDKDIE 120
Db 61 GRTLLKPDVDMGVASMIHEVGIEAMFPDGTGLVTHTPIEANGKLVPGEL-FLKNEDIT 119
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHHFFFEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
Db 120 INEGKAVSVKXNVGDRPVQIGSHHFFFEVNRCLDFDREKTFGKRLDIASGTAVRFBPG 179
QY 181 QTRKVLIPLGSKKVTGMNGLVNNIADERHKKHAKDKAKSHGF 224
Db 180 BEKSVELIDIGGNRRIFGFNALVDROADNESKIALHRAKERGF 223

RESULT 8
AAR04579

ID AAR04579 standard; protein; 238 AA.

XX AAR04579;

XX 27-AUG-2003 (revised)

DT 25-MAR-2003 (revised)

DT 21-SEP-1990 (first entry)

XX Part of protein with urease activity.

XX Urease; probe.

XX Helicobacter pylori.

XX WO9004030-A.

XX 19-APR-1990.

XX 06-OCT-1988; 88FR-00013135.

XX 06-OCT-1988; 88FR-00013135.

XX (INSP) INST PASTEUR.

PA (INRM) INSERM INST NAT SANTS & RECH MED.

XX Labigne A;

XX WPI; 1990-147844/19.

DR N-PSDB; AAR04328.

XX New nucleotide sequences encoding Campylobacter pylori-ureaseants - and
PT derived vectors, transformants, protein, antibodies and probes, useful in
PT diagnosis, treatment and prevention of infections.

XX Claim 11; Page 34; 47pp; French.

PS The protein can be used for the prodn. of antibodies and to prepare

CC vaccines for the prevention /treatment of C. pylori infections. See also

CC AAR04578-81. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 27

CC -AUG-2003 to correct OS field.)

XX Sequence 238 AA;

Query Match

Best Local Similarity 51.7%; Score 610.5; DB 2; Length 238;

Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;

QY 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIYISAHIMDEARRGKKTVAOLMEE 60

Db 1 MKLTPKELDKMLHYAGELARRKKEGKILNYYVEAVALISAHIMEEARAGKKTAAELMQE 60

QY 61 CMHFLKDEVMFGVGNMVPDLGVEATFPDGTGLVTNNWPIEPDEHFKAGVEKFGCDKDIE 120

Db 61 GRTLLKPDVDMGVASMIHEVGIEAMFPDGTGLVTHTPIEANGKLVPGEL-FLKNEDIT 119

QY 121 LNAGKEVTELEVTNEGPKSLHVGSHHFFFEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
Db 120 INEGKAVSVKXNVGDRPVQIGSHHFFFEVNRCLDFDREKTFGKRLDIASGTAVRFBPG 179
QY 181 QTRKVLIPLGSKKVTGMNGLVNNIADERHKKHAKDKAKSHGF 224
Db 180 BEKSVELIDIGGNRRIFGFNALVDROADNESKIALHRAKERGF 223

RESULT 9
AAW07193

ID AAW07193 standard; protein; 238 AA.

XX AAW07193;

DT 16-OCT-2003 (revised)

DT 11-FEB-1997 (first entry)

XX H. pylori urease A subunit.

XX Urease; ureA gene; ureB gene; vaccine.

XX Helicobacter pylori; strain CPM630.

XX WO9633732-A1.

XX 31-OCT-1996.

XX 25-APR-1996; 96WO-US005800.

XX 28-APR-1995; 95US-00431041.

PR 06-DEC-1995; 95US-00568122.

XX (ORAV-) ORAVAX INC.

XX Lee CK, Monath TP, Ackerman SK, Thomas WD, Sonan G, Kleanthous H;

PI Weltzin RA, Pappo J, Ermak T, Guirakhoo F, Bhagat H, Sussman I;

XX WPI; 1996-497373/49.

DR N-PSDB; AAT44351.

XX Vaccine for inducing mucosal response to Helicobacter - contg. multimeric
PT urease complex and pref. an antibiotic, anti-secretory agent or bismuth
PT salt.

XX Disclosure; Page 70-71; 98pp; English.

XX Urease A (AAW07193) and B (AAW07194) subunits of Helicobacter pylori are
CC encoded by the ureA + ureB gene locus of clinical isolate CPM630.

CC Vectors, e.g. pORV214 (see also AAT44351), carrying the gene locus can be
CC used to transform host cells for the large-scale prodn. of recombinant,
CC enzymatically inactive, multimeric urease. The urease complex induces a
CC mucosal immune response that can treat or prevent Helicobacter, esp. H.

CC pylori, gastroduodenal infection. (Updated on 16-OCT-2003 to standardise
CC OS field)

XX Sequence 238 AA;

Query Match

Best Local Similarity 51.7%; Score 610.5; DB 2; Length 238;

Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;

QY 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIYISAHIMDEARRGKKTVAOLMEE 60

Db 1 MKLTPKELDKMLHYAGELARRKKEGKILNYYVEAVALISAHIMEEARAGKKTAAELMQE 60

QY 61 CMHFLKDEVMFGVGNMVPDLGVEATFPDGTGLVTNNWPIEPDEHFKAGVEKFGCDKDIE 120

Db 61 GRTLLKPDVDMGVASMIHEVGIEAMFPDGTGLVTHTPIEANGKLVPGEL-FLKNEDIT 119

QY 121 LNAGKEVTELEVTNEGPKSLHVGSHHFFFEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180

Db 120 INEGKAVSVKXNVGDRPVQIGSHHFFFEVNRCLDFDREKTFGKRLDIASGTAVRFBPG 179

| | | |
|--|--|--|
| Query Match | 51.7%; Score 610.5; DB 5; Length 238; | CC atrophic gastritis, peptic ulcer diseases (e.g. gastric or duodenal |
| Best Local Similarity | 52.7%; Pred. No. 1.2e-55; | CC ulcers), and gastric carcinoma. Attenuated Salmonella vector pHUR3 was |
| Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1; | | CC constructed expressing UreA and UreB and including an E. coli AmpR gene. |
| QY 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEATAYISAHIMDEARRGKKTVAQLMEE 60 | | CC The present sequence represents a protein or peptide encoded by the pHUR3 |
| DB 1 MKLTPKELDKMLHYAGELAKRKEGKIKLNYVEAVALISAHIMEARAGKKTAAELMQE 60 | | CC plasmid. NOTE: pHUR3 is represented by both AEB55112 and AEB55130, the |
| QY 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTGLVTVMPIEPDEHFHFKAGEVKFGCDKDIE 120 | | CC peptides expressed by pHUR3 are shown in the sequence listing to be split |
| DB 61 GRTLKPDVDMGVASMIHEVGIEAMPDGTGLVTVHTPIEANGKLVPGEL-FLKNEDIT 119 | | CC between AEB55112 and AEB55130, yet Figure 4 shows all the peptides being |
| QY 121 LNAQKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180 | | CC expressed by the one sequence, therefore all encoded peptides are cross- |
| DB 120 INEGKAVSVKVNVDGDRPVQIGSHFHFVEVNRCLDFDREKTFGKRLDIASGTAVRFPFG 179 | | CC referenced to both sequences. |
| QY 181 QTRKVQLIPLGSKKVKVGMNGLVNNIADERHKHAKLDKAKSHGF 224 | | XX |
| DB 180 EEKSVELIDIGNRRIFGFNALVDROADNESKIALHRAKERGF 223 | | SQ Sequence 245 AA; |
| RESULT 12 | | |
| ID AEB55115 standard; protein; 245 AA. | Query Match 51.7%; Score 610.5; DB 4; Length 245; | |
| XX AC AEB55115; | Best Local Similarity 52.7%; Pred. No. 1.3e-55; | |
| XX DT 22-SEP-2005 (first entry) | Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1; | |
| DE DE Helicobacter pylori Urase A protein fragment SEQ ID 4. | | |
| KW pHUR3; vaccine; helicobacter pylori infection; urease; | QY 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEATAYISAHIMDEARRGKKTVAQLMEE 60 | |
| KW immune stimulation; antibacterial; Immunostimulant; antiulcer; | DB 1 MKLTPKELDKMLHYAGELAKRKEGKIKLNYVEAVALISAHIMEARAGKKTAAELMQE 67 | |
| KW cytostatic; antiinflammatory; gastrointestinal-gen.; Gastritis; | QY 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTGLVTVMPIEPDEHFHFKAGEVKFGCDKDIE 120 | |
| KW peptic ulcer; carcinoma. | DB 68 GRTLKPDVDMGVASMIHEVGIEAMPDGTGLVTVHTPIEANGKLVPGEL-FLKNEDIT 126 | |
| XX OS Helicobacter pylori. | QY 121 LNAQKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180 | |
| OS Salmonella typhimurium. | DB 127 INEGKAVSVKVNVDGDRPVQIGSHFHFVEVNRCLDFDREKTFGKRLDIASGTAVRFPFG 186 | |
| XX PN WO200132014-A2. | QY 181 QTRKVQLIPLGSKKVKVGMNGLVNNIADERHKHAKLDKAKSHGF 224 | |
| PD PD 10-MAY-2001. | DB 187 EEKSVELIDIGNRRIFGFNALVDROADNESKIALHRAKERGF 230 | |
| PP 01-NOV-2000; 2000WO-US030191. | RESULT 13 | |
| XX 01-NOV-1999; 99US-00431705. | ID ADS09178 standard; protein; 238 AA. | |
| XX (ORAV-) ORAVAX INC. | XX AC ADS09178; | |
| PI Kleanthous H, Londono-Arcilla P, Freeman D; | XX DT 16-DEC-2004 (first entry) | |
| DR WPI; 2001-343379/36. | DE DE H. pylori urease alpha beta subunit. | |
| DR N-PSDB; AEB55112, AEB55130. | KW Urease; alpha subunit; UreB; gastric ulcer; peptic ulcer; gastritis; | |
| PT Inducing an immune response against Helicobacter in mammals, useful for | XX stomach cancer; vaccine; antibody; immune reaction. | |
| PT treating Helicobacter induced gastroduodenal diseases. | XX OS Helicobacter pylori. | |
| XX Disclosure; SEQ ID NO 4; 63pp; English. | PN JP2004261080-A. | |
| CC The invention relates to inducing an immune response against Helicobacter | PD 24-SEP-2004. | |
| CC in a mammal, comprising mucosally administering to the mammal an | XX 28-FEB-2003; 2003JP-00054654. | |
| CC attenuated Salmonella vector containing a nucleic acid molecule encoding | XX 28-FEB-2003; 2003JP-00054654. | |
| CC a Helicobacter antigen, and parenterally administering to the mammal a | XX (DOKU-) DOKURITSU GYOSEI HOJIN KAGAKU GIJUTSU SH. | |
| CC Helicobacter antigen. Also included is an attenuated Salmonella vector | PA (FUKU-) FUKUYAMA RINSHO KENSA CENT KK. | |
| CC comprising a nucleic acid molecule encoding a Helicobacter antigen. The | XX WPI; 2004-665475/65. | |
| CC Helicobacter antigen is a urease, a urease subunit, or its immunogenic | PT Protein having specific region of beta-subunit urease of Helicobacter | |
| CC fragment (encoded by the ureA and ureB genes). The mammal is at risk of | PT pylori, which generates antibodies that suppress proliferation and urease | |
| CC having but does not have Helicobacter infection or has a Helicobacter | PT activity of bacteria, useful for preventing Helicobacter pylori | |
| CC infection. The attenuated Salmonella vector further comprises an htrA | XX infection. | |
| CC nirB promoter. The vector can be used in inducing an immune response | PS Disclosure; SEQ ID NO 3; 26pp; Japanese. | |
| CC against Helicobacter in a mammal. The vector can be used to treat | XX The invention relates to a protein having a specific region of primary | |
| CC Helicobacter infection. The vector and the method can be used to treat | CC structure of the beta-subunit Helicobacter pylori urease, appearing as | |
| CC Helicobacter induced gastroduodenal diseases, including acute, chronic or | CC ADS09176 (amino acids 201-338 of the beta subunit) or a sequence which | |
| | CC generates antibodies that suppresses proliferation of H. pylori and which | |
| | CC suppresses urease activity of H. pylori by an immune reaction. Also | |

CC included are a gene which encodes the urease fragment, a vaccine against
 CC H. pylori comprising the urease fragment, a transformed cell transformed
 CC by introducing the gene, an antibody which recognises the urease fragment
 CC and transforming a plant by introducing the gene into the plant and
 CC making it express. The urease fragment is useful for producing blood
 CC serum containing antibodies which suppresses proliferation of H. pylori
 CC and which suppresses urease activity of H. pylori which involves
 CC immunising a living organism (e.g. a human, cow or hen) with the urease
 CC fragment. The fragment may be purified from cow's milk or hen eggs from
 CC transgenic animals expressing the fragment. H. pylori is thought to be a
 CC cause of gastric and peptic ulcers, gastritis and may be implicated in
 CC stomach cancer. The urease enzyme is thought to be responsible for
 CC maintaining the bacterium in the acid conditions of the stomach. The
 CC present sequence is the H. pylori urease alpha subunit.

XX Sequence 238 AA;

Query Match 51.4%; Score 607.5; DB 8; Length 238;
 Best Local Similarity 52.2%; Pred. No. 2.6e-55;
 Matches 117; Conservative 44; Mismatches 62; Indels 1; Gaps 1;
 QY 1 VKLTPEQEKFLLYAGVAVRKAEGLKLNQPEALVISAHIMDEARRGKKTVAQLMEE 60
 DB 1 MKLTPKELDKMLHYAGELAKRKEGKIKNTVEAVALISAHIMEEARAGKKTAAELMQE 60
 QY 61 CMHFLKKDEVMGVMGNNVDPDLGVEATFPDGTKLVTNNWPIEDPDEHFKAGEVKGCDKDIE 120
 DB 61 GRTLLKPDVMDGVASMIHEVGIEAMFPDGTKLVTVHTPIEANGKLVPGEL-FLKNEDIT 119
 QY 121 LNAGKEVTELVNTEGPKSLHVGSHFFHFFANKALKFDREKAYGKRLDIPSGNTLIRIGAG 180
 DB 120 INEGKAVSVKVNVDGPRVQIGSHFFHFFVNRCLDFDREKTFGKRLDIAAGTAVRFPFG 179
 QY 181 QTRKVLQPLGSKKVIKGNGLVNNIADERHKHKALDKAKSHGF 224
 DB 180 EEKVELIDIGNRRIFGFGNALVDQADNESKKIALHRAKERGF 223

RESULT 14

ADM28643
 ID ADM28643 standard; protein; 238 AA.

AC ADM28643;

XX 20-MAY-2004 (first entry)

DT Helicobacter pylori urease alpha subunit protein SeqID 2.

DE Helicobacter pylori urease alpha subunit protein SeqID 2.

KW immunogenic; urease; vaccine; passive immunisation; diagnostic;

KW antibacterial; antibody.

XX Helicobacter pylori.

OS Helicobacter pylori.

PN JP2004041084-A.

XX 12-FEB-2004.

PF 11-JUL-2002; 2002JP-00203221.

XX 11-JUL-2002; 2002JP-00203221.

PR (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.

XX WPI; 2004-162021/16.

DR New peptide useful as a vaccine against Helicobacter pylori induces an

PT antibody response against the bacterial urease.

XX Disclosure; SEQ ID NO 2; 19pp; Japanese.

XX This invention relates to a novel immunogenic peptide that induces

CC production of antibodies against the Helicobacter pylori (H. pylori)

CC urease protein. Specifically, it refers to a peptide antigen capable of

CC inducing an antibody response, such that it can be used to develop a
 CC vaccine against H. pylori. The present invention describes generating
 CC monoclonal antibodies against the urease using hybridoma techniques that
 CC are useful for passive immunisation and diagnostic techniques.
 CC Furthermore, the antibacterial peptide can be used to immunise an animal
 CC in order to produce antibodies that are present within a functional food
 CC such as milk or eggs. This polypeptide sequence is the H. pylori urease
 CC alpha subunit protein of the invention.

XX Sequence 238 AA;

Query Match 51.1%; Score 603.5; DB 8; Length 238;
 Best Local Similarity 52.2%; Pred. No. 6.8e-55;
 Matches 117; Conservative 43; Mismatches 63; Indels 1; Gaps 1;

QY 1 VKLTPEQEKFLLYAGVAVRKAEGLKLNQPEALVISAHIMDEARRGKKTVAQLMEE 60
 DB 1 MKLTPKELDKMLHYAGELAKRKEGKIKNTVEAVALISAHIMEEARAGKKTAAELMQE 60
 QY 61 CMHFLKKDEVMGVMGNNVDPDLGVEATFPDGTKLVTNNWPIEDPDEHFKAGEVKGCDKDIE 120
 DB 61 GRTLLKPDVMDGVASMIHEVGIEAMFPDGTKLVTVHTPIEANGKLVPGEL-FLKNEDIT 119
 QY 121 LNAGKEVTELVNTEGPKSLHVGSHFFHFFANKALKFDREKAYGKRLDIPSGNTLIRIGAG 180
 DB 120 INEGKAVSVKVNVDGPRVQIGSHFFHFFVNRCLDFDREKTFGKRLDIAAGTAVRFPFG 179
 QY 181 QTRKVLQPLGSKKVIKGNGLVNNIADERHKHKALDKAKSHGF 224
 DB 180 EEKVELIDIGNRRIFGFGNALVDQADNESKKIALHRAKERGF 223

RESULT 15

AAR67375
 ID AAR67375 standard; protein; 238 AA.

AC AAR67375;

XX 25-MAR-2003 (revised)

DT 22-JUN-1995 (first entry)

DE H. pylori ureA urease.

KW Urease; ureA gene; immunogen; vaccine; diagnostic; Helicobacter felis.

XX Proteus mirabilis.

XX WO9426901-A1.

PD 24-NOV-1994.

XX 19-MAY-1994; 94WO-EP001625.

XX 19-MAY-1993; 93EP-00401309.

PR 19-NOV-1993; 93WO-EP003259.

XX (INSP) INST PASTEUR.

PA (INRM) INST NAT SANTE & RECH MEDICALE.

XX Labigne A, Suerbaum S, Ferrero R, Thiberge J;

XX WPI; 1995-006797/01.

PT DNA from Helicobacter pylori and Helicobacter felis - used to develop

PT prods. for detection, treatment and prevention of Helicobacter infection.

XX Disclosure; Fig 4i-iii; 168pp; English.

XX The sequence of the Helicobacter felis urease ureA gene product (given in

CC AAR67371) was compared to ureases of Helicobacter pylori (AAR67375),

CC Proteus mirabilis (AAR67376) and jack bean urease (AAR67377) and regions

CC of homology were identified. (Updated on 25-MAR-2003 to correct PN

CC field.)

```

XX
SQ      Sequence 238 AA;
      Query Match      51.0%; Score 602.5; DB 2; Length 238;
      Best Local Similarity 52.2%; Pred. No. 8.7e-55;
      Matches 117; Conservative 43; Mismatches 63; Indels 1; Gaps 1;

QY      1 VKLTPKEQKFLLYAGEVARKKAEGKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
      :||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1 MKLTPKELDKLMLHYAGELAKRKEGIKLNYVEAVALISAHIMEEERAGKTADELMOE 60

QY      61 CMHFLKKDEVMPGVGNMVPDLGVEATPPDGTCLVTNNWPIEPDEHFKAGEVKFGCDKDIE 120
      || :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db      61 GRTLLKPDVVMDGVASMIHEVGIEAMPDPDGTCLVTHTPIEANGKLVPGEL-FLKNEDIT 119

QY      121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKEDREKAYCKRLDIPSGNTLRICAG 180
      :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db      120 INEGKAVSVKKNVGDPRPVOIGSHFHFVEVNRCLDPDREKTFHRLUDIASGTAVRFPFG 179

QY      181 QTRKVOLIPLGSKKQVIGMNLVNNIADERRHKIKALDKAKSHGF 224
      : :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db      180 EEKSVELLIDIGNRRIFGFNALVDRQADNESKKIALHRAKERGF 223
```

Search completed: November 28, 2005, 08:14:34
Job time : 76.7204 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 28, 2005, 08:11:38 ; Search time 63.7582 Seconds
(without alignments)
1481.056 Million cell updates/sec

Title: US-09-904-994B-2
Perfect score: 1181
Sequence: 1 VKLTPKEQKFLLYAGEVA.....ADERHKHAKLDKAKSHGFIK 226

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA_Main:
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pap:
2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pap:
3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pap:
4: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pap:
5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pap:
6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|----------------------|
| 1 | 1181 | 100.0 | 226 | 3 | US-09-904-994B-2 |
| 2 | 1178 | 99.7 | 226 | 3 | US-09-904-994B-14 |
| 3 | 1174 | 99.4 | 226 | 3 | US-09-904-994B-5 |
| 4 | 1174 | 99.4 | 226 | 3 | US-09-904-994B-11 |
| 5 | 1170 | 99.1 | 226 | 3 | US-09-904-994B-8 |
| 6 | 652 | 55.2 | 234 | 4 | US-10-639-273-3 |
| 7 | 622 | 52.7 | 234 | 4 | US-10-639-273-34 |
| 8 | 613.5 | 51.9 | 238 | 3 | US-09-915-242-11468 |
| 9 | 613.5 | 51.9 | 238 | 4 | US-10-335-977-8629 |
| 10 | 610.5 | 51.7 | 238 | 3 | US-09-915-242-11286 |
| 11 | 594 | 50.3 | 237 | 4 | US-10-639-273-36 |
| 12 | 576.5 | 48.8 | 228 | 4 | US-10-335-977-8628 |
| 13 | 573.5 | 48.6 | 228 | 4 | US-10-012-819-154 |
| 14 | 566.5 | 48.0 | 237 | 4 | US-10-335-977-8629 |
| 15 | 540.5 | 45.8 | 225 | 4 | US-10-639-273-37 |
| 16 | 502.5 | 42.5 | 811 | 4 | US-10-282-122A-70574 |
| 17 | 497.5 | 42.1 | 207 | 4 | US-10-282-122A-70447 |
| 18 | 493 | 41.7 | 779 | 4 | US-10-282-122A-69134 |
| 19 | 489.5 | 41.4 | 207 | 3 | US-09-915-242-5448 |
| 20 | 477 | 40.4 | 226 | 4 | US-10-282-122A-48209 |
| 21 | 470.5 | 39.8 | 209 | 4 | US-10-282-122A-68251 |
| 22 | 470 | 39.8 | 261 | 4 | US-10-282-122A-55507 |
| 23 | 469.5 | 39.8 | 207 | 4 | US-10-282-122A-45038 |
| 24 | 455.5 | 38.6 | 228 | 4 | US-10-282-122A-50077 |
| 25 | 439.5 | 37.2 | 840 | 4 | US-10-621-833-7 |
| 26 | 439.5 | 37.2 | 840 | 5 | US-10-731-877-1 |
| 27 | 439.5 | 37.2 | 840 | 6 | US-11-046-271-1 |

| | | | | | | |
|----|-------|------|-----|---|----------------------|-------------------|
| 28 | 426 | 36.1 | 843 | 4 | US-10-437-963-124057 | Sequence 124057, |
| 29 | 424.5 | 35.9 | 171 | 4 | US-10-335-977-8627 | Sequence 8627, Ap |
| 30 | 384.5 | 32.6 | 227 | 4 | US-10-156-761-10253 | Sequence 10253, A |
| 31 | 381.5 | 32.3 | 837 | 4 | US-10-424-599-254635 | Sequence 254635, |
| 32 | 368.5 | 31.2 | 837 | 4 | US-10-418-962-2 | Sequence 2, Appli |
| 33 | 291 | 24.6 | 227 | 4 | US-10-425-115-201340 | Sequence 201340, |
| 34 | 281.5 | 23.8 | 100 | 3 | US-09-815-242-11745 | Sequence 11745, A |
| 35 | 281.5 | 23.8 | 100 | 4 | US-10-156-761-14633 | Sequence 14633, A |
| 36 | 281.5 | 23.8 | 100 | 4 | US-10-282-122A-60250 | Sequence 60250, A |
| 37 | 279.5 | 23.7 | 121 | 4 | US-10-282-122A-49772 | Sequence 49772, A |
| 38 | 278.5 | 23.6 | 100 | 4 | US-10-282-122A-62539 | Sequence 62539, A |
| 39 | 278.5 | 23.6 | 100 | 4 | US-10-282-122A-64616 | Sequence 64616, A |
| 40 | 277.5 | 23.5 | 100 | 4 | US-10-282-122A-56822 | Sequence 56822, A |
| 41 | 277.5 | 23.5 | 100 | 4 | US-10-282-122A-56846 | Sequence 56846, A |
| 42 | 267.5 | 22.7 | 100 | 3 | US-09-815-242-11047 | Sequence 11047, A |
| 43 | 267.5 | 22.7 | 100 | 4 | US-10-282-122A-58211 | Sequence 58211, A |
| 44 | 267.5 | 22.7 | 100 | 4 | US-10-282-122A-69398 | Sequence 69398, A |
| 45 | 265.5 | 22.5 | 100 | 4 | US-10-282-122A-44442 | Sequence 44442, A |

ALIGNMENTS

RESULT 1

US-09-904-994B-2
; Sequence 2, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO NO. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-2

Query Match 100.0%; Score 1181; DB 3; Length 226;
Best Local Similarity 100.0%; Pred. No. 3.4e-114;
Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| Qy | 1 | VKLTPKEQKFLLYAGEVARKKAEGLKLNQPEAIYISAHIMDEARRGKKTVAQLMEE | 60 |
| Db | 1 | VKLTPKEQKFLLYAGEVARKKAEGLKLNQPEAIYISAHIMDEARRGKKTVAQLMEE | 60 |
| Qy | 61 | CMHFLKKDEVMPGVGNMVPDLGVETATPDGKLVTVNWPIDPDEHFHFKAGEVFGCDKDI | 120 |
| Db | 61 | CMHFLKKDEVMPGVGNMVPDLGVETATPDGKLVTVNWPIDPDEHFHFKAGEVFGCDKDI | 120 |
| Qy | 121 | LNAGKEVTELEVTNEGPKSLHVGSHFFHFEANKALKFDPREKAYKRLDIPSGNTLRIGAG | 180 |
| Db | 121 | LNAGKEVTELEVTNEGPKSLHVGSHFFHFEANKALKFDPREKAYKRLDIPSGNTLRIGAG | 180 |
| Qy | 181 | QTRKVLIPLGSKKVGCMGNLVNNIADERHKHAKLDKAKSHGFIK | 226 |
| Db | 181 | QTRKVLIPLGSKKVGCMGNLVNNIADERHKHAKLDKAKSHGFIK | 226 |

RESULT 2

US-09-904-994B-14
; Sequence 14, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO NO. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-14

Query Match      99.7%; Score 1178; DB 3; Length 226;
Best Local Similarity 99.6%; Pred. No. 7.1e-114;
Matches 225; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
   |||||
Db 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAELMEE 60
   |||||
QY 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTCLVTNNWPIEPDEHFKAGEVKGCDKDIE 120
   |||||
Db 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTCLVTNNWPIEPDEHFKAGEVKGCDKDIE 120
   |||||
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFFEANKALFDRKAYGKRLDIPSGNTLRIGAG 180
   |||||
Db 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFFEANKALFDRKAYGKRLDIPSGNTLRIGAG 180
   |||||
QY 181 QTRKVQLIPLGGSKKVI GNMGLVNNIADERRHKH KALDKAKSHGFIK 226
   |||||
Db 181 QTRKVQLIPLGGSKKVI GNMGLVNNIADERRHKH KALDKAKSHGFIK 226
   |||||

RESULT 3
US-09-904-994B-5
; Sequence 5, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-5

Query Match      99.4%; Score 1174; DB 3; Length 226;
Best Local Similarity 99.1%; Pred. No. 1.8e-113;
Matches 224; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
   |||||
Db 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAELMEE 60
   |||||
QY 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTCLVTNNWPIEPDEHFKAGEVKGCDKDIE 120
   |||||
Db 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTCLVTNNWPIEPDEHFKAGEVKGCDKDIE 120
   |||||
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFFEANKALFDRKAYGKRLDIPSGNTLRIGAG 180
   |||||
Db 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFFEANKALFDRKAYGKRLDIPSGNTLRIGAG 180
   |||||
QY 181 QTRKVQLIPLGGSKKVI GNMGLVNNIADERRHKH KALDKAKSHGFIK 226
   |||||
Db 181 QTRKVQLIPLGGSKKVI GNMGLVNNIADERRHKH KALDKAKSHGFIK 226
   |||||

RESULT 4
US-09-904-994B-11
; Sequence 11, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-11

Query Match      99.4%; Score 1174; DB 3; Length 226;
Best Local Similarity 99.1%; Pred. No. 1.8e-113;
Matches 224; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
   |||||
Db 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAELMEE 60
   |||||
QY 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTCLVTNNWPIEPDEHFKAGEVKGCDKDIE 120
   |||||
Db 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTCLVTNNWPIEPDEHFKAGEVKGCDKDIE 120
   |||||
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFFEANKALFDRKAYGKRLDIPSGNTLRIGAG 180
   |||||
Db 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFFEANKALFDRKAYGKRLDIPSGNTLRIGAG 180
   |||||
QY 181 QTRKVQLIPLGGSKKVI GNMGLVNNIADERRHKH KALDKAKSHGFIK 226
   |||||
Db 181 QTRKVQLIPLGGSKKVI GNMGLVNNIADERRHKH KALDKAKSHGFIK 226
   |||||

RESULT 5
US-09-904-994B-8
; Sequence 8, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-8

Query Match      99.1%; Score 1170; DB 3; Length 226;
Best Local Similarity 98.7%; Pred. No. 4.8e-113;
Matches 223; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
   |||||
Db 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAELMEE 60
   |||||
QY 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTCLVTNNWPIEPDEHFKAGEVKGCDKDIE 120
   |||||
Db 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTCLVTNNWPIEPDEHFKAGEVKGCDKDIE 120
   |||||
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFFEANKALFDRKAYGKRLDIPSGNTLRIGAG 180
   |||||
Db 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFFEANKALFDRKAYGKRLDIPSGNTLRIGAG 180
   |||||
QY 181 QTRKVQLIPLGGSKKVI GNMGLVNNIADERRHKH KALDKAKSHGFIK 226
   |||||
Db 181 QTRKVQLIPLGGSKKVI GNMGLVNNIADERRHKH KALDKAKSHGFIK 226
   |||||

RESULT 6
US-09-904-994B-8
; Sequence 8, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-8

Query Match      99.1%; Score 1170; DB 3; Length 226;
Best Local Similarity 98.7%; Pred. No. 4.8e-113;
Matches 223; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
   |||||
Db 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAELMEE 60
   |||||
QY 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTCLVTNNWPIEPDEHFKAGEVKGCDKDIE 120
   |||||
Db 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTCLVTNNWPIEPDEHFKAGEVKGCDKDIE 120
   |||||
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFFEANKALFDRKAYGKRLDIPSGNTLRIGAG 180
   |||||
Db 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFFEANKALFDRKAYGKRLDIPSGNTLRIGAG 180
   |||||
QY 181 QTRKVQLIPLGGSKKVI GNMGLVNNIADERRHKH KALDKAKSHGFIK 226
   |||||
Db 181 QTRKVQLIPLGGSKKVI GNMGLVNNIADERRHKH KALDKAKSHGFIK 226
   |||||
```

```
US-10-639-273-3
; Sequence 3, Application US/10639273
; Publication No. US20040142343A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; APPLICANT: Simpson, Kenneth W
; APPLICANT: Zhu, Jiaqian
; TITLE OF INVENTION: HELICOBACTER BIZZOZERONII UREASE GENES AND THEIR USES
; TITLE OF INVENTION: IN DIAGNOSTIC AND TREATMENT METHODS
; FILE REFERENCE: 19603/3881
; CURRENT APPLICATION NUMBER: US/10/639,273
; CURRENT FILING DATE: 2003-08-12
; PRIOR APPLICATION NUMBER: 60/404,337
; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Helicobacter bizzozeronii
US-10-639-273-3

Query Match 55.2%; Score 652; DB 4; Length 234;
Best Local Similarity 57.3%; Pred. No. 3.9e-59;
Matches 129; Conservative 39; Mismatches 55; Indels 2; Gaps 2;

QY 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEATAYISAHIMDEARRGKKTVAOLMEE 60
DB 1 MKLTPKELDKMLHYAGELARKKAEGLKLNQPEATAYISAHIMDEARRGKKTVAOLMEE 60

QY 61 CMHFLKKDEVMGPNVDPDLGVEATFPDGTGKLVTVNWPIEP-DEHFKAGEVKFGCDKDI 119
DB 61 GRTLKADDDVMGVAHMIHEVIEAGFPDGTGKLVTVHTFVIEAGSKLAPGEVILK-NEDI 119

QY 120 ELNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDRKAYGKRLDIPSGNTLRIGA 179
DB 120 TLMAGKQATTLVHNKGDPRVQVGSFHFFEVNKLLEFDRKAYGKRLDIASGTAVRPEP 179

QY 180 GQTRKQVLIPLGGSKKVGIMGNLVNNIADDERHKHKALDKAKSHGF 224
DB 180 GEKTVELIDIGGNRIYGFNSLVDRQADTDGKKLAKRAKEHGF 224

RESULT 7
US-10-639-273-34
; Sequence 34, Application US/10639273
; Publication No. US20040142343A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; APPLICANT: Simpson, Kenneth W
; APPLICANT: Zhu, Jiaqian
; TITLE OF INVENTION: HELICOBACTER BIZZOZERONII UREASE GENES AND THEIR USES
; TITLE OF INVENTION: IN DIAGNOSTIC AND TREATMENT METHODS
; FILE REFERENCE: 19603/3881
; CURRENT APPLICATION NUMBER: US/10/639,273
; CURRENT FILING DATE: 2003-08-12
; PRIOR APPLICATION NUMBER: 60/404,337
; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Helicobacter heilmannii
US-10-639-273-34

Query Match 52.7%; Score 622; DB 4; Length 234;
Best Local Similarity 54.7%; Pred. No. 5.2e-56;
Matches 123; Conservative 41; Mismatches 59; Indels 2; Gaps 2;

QY 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEATAYISAHIMDEARRGKKTVAOLMEE 60
DB 1 MKLTPKELDKMLHYAGELARKKAEGLKLNQPEATAYISAHIMDEARRGKKTVAOLMEE 60

US-09-815-242-11468
; Sequence 11468, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11468
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-815-242-11468

Query Match 51.9%; Score 613.5; DB 3; Length 238;
Best Local Similarity 53.1%; Pred. No. 4.1e-55;
Matches 119; Conservative 42; Mismatches 62; Indels 1; Gaps 1;

QY 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAIYISAHIMDEARRGKKTVAOLMEE 60
DB 1 MKLTPKELDKMLHYAGELARKKAEGLKLNQPEAIAIYISAHIMDEARRGKKTVAOLMEE 60

QY 61 CMHFLKKDEVMGPNVDPDLGVEATFPDGTGKLVTVNWPIEPDEHFKAGEVKFGCDKDI 120
DB 61 GRTLKADDDVMGVAHMIHEVIEAGFPDGTGKLVTVHTFVIEAGSKLAPGEVILK-NEDI 119

QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDRKAYGKRLDIPSGNTLRIGA 180
DB 120 INEGKKAIVSVKXNVGNDRPVQVGSFHFFEVNKLLEFDRKAYGKRLDIASGTAVRPEP 179

QY 181 QTRKQVLIPLGGSKKVGIMGNLVNNIADDERHKHKALDKAKSHGF 224
DB 180 ESKVELIDIGGNRIYGFNSLVDRQADTDGKKLAKRAKEHGF 223
```

RESULT 9
US-10-335-977-8629
; Sequence 8629, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 8629:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...238
; SEQUENCE DESCRIPTION: SEQ ID NO: 8629:
US-10-335-977-8629

Query Match 51.9%; Score 613.5; DB 4; Length 238;
Best Local Similarity 53.1%; Pred. No. 4.1e-55;
Matches 119; Conservative 42; Mismatches 62; Indels 1; Gaps 1;
QY 1 VKLTPKEQKFLYYAGVARKKAEGLKLNQPEAIYISAHIMDEARRGKKTVAOLMEE 60
Db 1 MKLTPKELDKMLHYAGELARRKKGKILNYYVEAVALLISAHIMEEARAGKKTAAELMQE 60
QY 61 CMHFLKKDEVMGPNVGNVDPDLGVEATFPDGTKLVTVMPIEDPHFKAQVKGCDKDIE 120
Db 61 GRTLLXPDDVMDGVASMIHEVGIEAMFPDGTKLVTVHTPIEANGKLVPGEL-FLKNEDIT 119
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
Db 120 INEGKAVSVKVNVDGDRPVQIGSHFHFPEVNRCLDFDREKTGKRLDIASGTAVRFPFG 179
QY 181 QTRKVVLIPLGGSKKVGIMGNLVNNTADERHKKHAKDKAKSHGF 224
Db 180 EEKSVELIDIGNRRIFGFGNALVDROADNESKKIALHRAKERGF 223

RESULT 10
US-09-815-242-11286
; Sequence 11286, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11286
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-815-242-11286
Query Match 51.7%; Score 610.5; DB 3; Length 238;
Best Local Similarity 52.7%; Pred. No. 8.4e-55;
Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;
QY 1 VKLTPKEQKFLYYAGVARKKAEGLKLNQPEAIYISAHIMDEARRGKKTVAOLMEE 60
Db 1 MKLTPKELDKMLHYAGELARRKKGKILNYYVEAVALLISAHIMEEARAGKKTAAELMQE 60
QY 61 CMHFLKKDEVMGPNVGNVDPDLGVEATFPDGTKLVTVMPIEDPHFKAQVKGCDKDIE 120
Db 61 GRTLLXPDDVMDGVASMIHEVGIEAMFPDGTKLVTVHTPIEANGKLVPGEL-FLKNEDIT 119
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
Db 120 INEGKAVSVKVNVDGDRPVQIGSHFHFPEVNRCLDFDREKTGKRLDIASGTAVRFPFG 179
QY 181 QTRKVVLIPLGGSKKVGIMGNLVNNTADERHKKHAKDKAKSHGF 224
Db 180 EEKSVELIDIGNRRIFGFGNALVDROADNESKKIALHRAKERGF 223
RESULT 11
US-10-639-273-36
; Sequence 36, Application US/10639273
; Publication No. US20040142343A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; APPLICANT: Simpson, Kenneth W
; APPLICANT: Zhu, Jiaqian
; TITLE OF INVENTION: HELICOBACTER BIZZOZERONII UREASE GENES AND THEIR USES
; TITLE OF INVENTION: IN DIAGNOSTIC AND TREATMENT METHODS
; FILE REFERENCE: 19603/3881
; CURRENT APPLICATION NUMBER: US/10/639,273

; CURRENT FILING DATE: 2003-08-12
; PRIOR APPLICATION NUMBER: 60/404,337
; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-10-639-273-36

Query Match 50.3%; Score 594; DB 4; Length 237;
Best Local Similarity 52.2%; Pred. No. 4.4e-53;
Matches 117; Conservative 43; Mismatches 62; Indels 2; Gaps 2;

Qy 1 VKLTPEQKQLLYAGVARRKKAAGLKLNQPEAIYISAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLTPKELDKMLHYAGELAKRKEGKIKLNYVEAVRLISAHIMEEARRGKKTAAELMQE 60

Qy 61 CMHFLKKDEVMPGVGNVVDLGEATFPDGTCLVTNNWPIEPDEHFKAGEVFGCDKDIE 120
Db 61 GRTLKPPDDVMDGVASMIHEVGIEAMFPDGTCLVTVHTPIEANGKLVPGEL-FLKNEDIT 119

Qy 121 LNAGKEVTELVNNEGPKSLHVGSHFFFEVNRCLDPDEKAYGKRLDIPSGNTLRIGAG 180
Db 120 INEGKAVSVKPPVGDPRPVQIGSHFFFEVNRCLDPDEKAYGKRLDIPSGNTLRIGAG 180

Qy 181 QTRKVOLIPLGSKVKVGNGLVNNIADERHKHKKALDKAKSHGF 224
Db 180 EKSVELIDIGGNRRIFGFNALVDROADNESK-IALHRAKERGF 222

RESULT 12
US-10-335-977-8628
; Sequence 8628, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandregouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 8628:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...228
; SEQUENCE DESCRIPTION: SEQ ID NO: 8628:
US-10-335-977-8628

Query Match 48.8%; Score 576.5; DB 4; Length 228;
Best Local Similarity 52.6%; Pred. No. 2.7e-51;
Matches 112; Conservative 40; Mismatches 60; Indels 1; Gaps 1;

Qy 12 LLYYAGEVARRKKAAGLKLNQPEAIYISAHIMDEARRGKKTVAQLMEECMHFLKKDEV 71
Db 2 MLHYAGELAKRKEGKIKLNYVEAVRLISAHIMEEARRGKKTAAELMQEGRTLKPPDDVM 61

Qy 72 PGVGNVVDLGEATFPDGTCLVTNNWPIEPDEHFKAGEVFGCDKDIELNAGKEVTELE 131
Db 62 DGVASMIHEVGIEAMFPDGTCLVTVHTPIEANGKLVPGEL-FLKNEDITINEGKAVSVK 120

Qy 132 VTNEGPKSLHVGSHFFFEVNRCLDPDEKAYGKRLDIPSGNTLRIGAGQTRKVOLIPLG 191
Db 121 VKWGDPRVQIGSHFFFEVNRCLDPDEKAYGKRLDIPSGNTLRIGAGQTRKVOLIPLG 180

Qy 192 GSKKVIQGNGLVNNIADERHKHKKALDKAKSHGF 224
Db 181 GNRRIFGFNLVDROADNESKIALHRAKERGF 213

RESULT 13
US-10-012-819-154
; Sequence 154, Application US/10012819
; Publication No. US20030017478A1
; GENERAL INFORMATION:
; APPLICANT: Legrain, Pierre
; APPLICANT: Selig, Luc
; APPLICANT: Rain, Jean-Christophe
; TITLE OF INVENTION: Collection of Prokaryotic DNA for Two-Hybrid System, Helicobacter
; FILE REFERENCE: B5053
; CURRENT APPLICATION NUMBER: US/10/012,819
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: EP 99401066.8
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 278
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 154
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-10-012-819-154

Query Match 48.6%; Score 573.5; DB 4; Length 228;
Best Local Similarity 52.1%; Pred. No. 5.6e-51;
Matches 111; Conservative 41; Mismatches 60; Indels 1; Gaps 1;

Qy 12 LLYYAGEVARRKKAAGLKLNQPEAIYISAHIMDEARRGKKTVAQLMEECMHFLKKDEV 71
Db 2 MLHYAGELAKRKEGKIKLNYVEAVRLISAHIMEEARRGKKTAAELMQEGRTLKPPDDVM 61

Qy 72 PGVGNVVDLGEATFPDGTCLVTNNWPIEPDEHFKAGEVFGCDKDIELNAGKEVTELE 131
Db 62 DGVASMIHEVGIEAMFPDGTCLVTVHTPIEANGKLVPGEL-FLKNEDITINEGKAVSVK 120

Qy 132 VTNEGPKSLHVGSHFFFEVNRCLDPDEKAYGKRLDIPSGNTLRIGAGQTRKVOLIPLG 191
Db 121 VKWGDPRVQIGSHFFFEVNRCLDPDEKAYGKRLDIPSGNTLRIGAGQTRKVOLIPLG 180

Qy 192 GSKKVIQGNGLVNNIADERHKHKKALDKAKSHGF 224
Db 181 GNRRIFGFNLVDROADNESKIALHRAKERGF 213

```

Db      1  MRUTPELDKMLHLYAGALAKSRERGIKUNYVESIALISWEIMELAREGNKSVAEMLMQQ 60
Qy      61  CMHFLKDDVYMPGVGNVDPDLGVEATEPFDGTKLVTVNWMPIDPDEHFKAAGEVKGCDKOIE 120
Db      61  GREILKADDEVGASVMVNEQVEVSPFDGTKLVTHNPIDEDNGKLTGPEYILK-DEDIT 119
Qy      121  LNAGKEVTELEVNEGKPSLHVSGSHFFFEANKALKPDREKAYGKRLLDIPSGNTURIGAG 180
Db      120  LNANKESISIKVTHKGDRPIQVSGSHFFFEYNALLEFDRAQAFGRKRLDIASGTSVRFEB 179
Qy      181  QTRKVLVLIPLGGSKVKVIGMGLVNNIADDERHKHKALDKAKSHGFI 225
Db      180  EEKNVLIDFGGQKXILGFNDLNTNAHINKENKEOCLANAAQKHFI 224

Search completed: November 28, 2005, 08:24:30
Job time : 64.7582 secs

```


| Result No. | Score | Match | Length | DB | ID | Description | Sequence 116, Appl |
|------------|-------|-------|--------|----|--------------------|--------------------|--------------------|
| | | | | | | | |
| 1 | 247 | 20.9 | 111 | 1 | US-10-793-626-116 | Sequence 116, Appl | Sequence 181, Appl |
| 2 | 212 | 18.0 | 106 | 1 | US-10-485-517-181 | Sequence 181, Appl | Sequence 2006, Ap |
| 3 | 84.5 | 7.2 | 247 | 1 | US-10-793-626-2006 | Sequence 2006, Ap | Sequence 108, App |
| 4 | 77.5 | 6.6 | 259 | 1 | US-10-510-386-108 | Sequence 108, App | Sequence 50, Appl |
| 5 | 76.5 | 6.5 | 428 | 1 | US-10-689-742-50 | Sequence 50, Appl | Sequence 1089, Ap |
| 6 | 73.5 | 6.2 | 419 | 1 | US-10-821-234-1089 | Sequence 1089, Ap | Sequence 6, Appl |
| 7 | 72 | 6.1 | 418 | 1 | US-10-858-730-6 | Sequence 6, Appl | Sequence 1248, Ap |
| 8 | 71 | 6.0 | 311 | 1 | US-10-793-626-1248 | Sequence 1248, Ap | Sequence 1460, Ap |
| 9 | 70 | 5.9 | 329 | 1 | US-10-793-626-1460 | Sequence 1460, Ap | Sequence 1608, Ap |
| 10 | 70 | 5.9 | 437 | 1 | US-10-821-234-1608 | Sequence 1608, Ap | Sequence 244, App |
| 11 | 69 | 5.8 | 501 | 1 | US-10-793-626-244 | Sequence 244, App | Sequence 3, Appl |
| 12 | 68.5 | 5.8 | 1992 | 7 | US-11-013-759-3 | Sequence 3, Appl | Sequence 13, Appl |
| 13 | 68.5 | 5.8 | 1992 | 7 | US-11-013-759-13 | Sequence 13, Appl | Sequence 4, Appl |
| 14 | 68.5 | 5.8 | 2047 | 7 | US-11-013-759-4 | Sequence 4, Appl | Sequence 7, Appl |
| 15 | 68.5 | 5.8 | 2047 | 7 | US-11-013-759-7 | Sequence 7, Appl | Sequence 65, Appl |
| 16 | 67 | 5.7 | 932 | 7 | US-11-017-550-65 | Sequence 65, Appl | Sequence 3184, Ap |
| 17 | 67 | 5.7 | 1279 | 7 | US-10-793-626-3188 | Sequence 3184, Ap | Sequence 2594, Ap |
| 18 | 66.5 | 5.6 | 567 | 1 | US-10-793-626-3184 | Sequence 2594, Ap | Sequence 462, App |
| 19 | 66.5 | 5.6 | 989 | 1 | US-10-793-626-2594 | Sequence 462, App | Sequence 1452, Ap |
| 20 | 66 | 5.6 | 600 | 1 | US-10-131-826A-462 | Sequence 1452, Ap | Sequence 51, Appl |
| 21 | 65.5 | 5.5 | 328 | 1 | US-10-821-234-1462 | Sequence 51, Appl | Sequence 2, Appl |
| 22 | 65.5 | 5.5 | 1015 | 1 | US-10-957-569-51 | Sequence 2, Appl | Sequence 172, App |
| 23 | 65.5 | 5.5 | 1168 | 1 | US-10-509-422-2 | Sequence 172, App | Sequence 2170, Ap |
| 24 | 65 | 5.5 | 397 | 1 | US-10-485-517-172 | Sequence 2170, Ap | Sequence 2170, Ap |
| 25 | 64 | 5.4 | 302 | 1 | US-10-793-626-2170 | Sequence 2170, Ap | |

```
; CURRENT FILING DATE: 2004-02-02
; PRIOR APPLICATION NUMBER: GB 0118825.9
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: GB 0200349.9
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 424
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 181
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-485-517-181

Query Match      18.0%; Score 212; DB 1; Length 106;
Best Local Similarity 46.9%; Pred. No. 2.4e-14;
Matches 38; Conservative 15; Mismatches 28; Indels 0; Gaps 0;

QY 138 KSLHVGSHHFFPEANKALKFDEKAYGKRLDIPSGNTLRIGAGQTRKVKQLIPLGGSKKVI 197
DB 2 RPIQVGSFHFHYEANAALDFEREMAYGKHLDPAGAAVRFEPGDKKEVOLVVEYAGRKIF 61
QY 198 GNGLVNNIADERHKHAKDK 218
DB 62 GFRGMVNGPIDESRVIRPTDE 82

RESULT 3
US-10-793-626-2006
; Sequence 2006, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2006
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-2006

Query Match      7.2%; Score 84.5; DB 1; Length 247;
Best Local Similarity 23.8%; Pred. No. 0.14;
Matches 48; Conservative 34; Mismatches 69; Indels 51; Gaps 12;

QY 44 MDEARRGK-----KTVAQLMEECHMFLKDKDEVMPGVGMVDPDLGVBEATFPDGTGL-- 93
DB 73 LEQARQGLAIHNDHMLHTIEQPREELSAYAPKVVTM----SINPKIRDVIGPGGKKINE 128
QY 94 ----VTVNWPIPDHFHFKAGEVFGCKDKDIELNAGKEVTELEVNTGPKSLHVGSHFHF 149
DB 129 IIDETGVKLDIEQD-----GTIFIGAVDQAMINRAKEIIE-DITREA---EVGVYH-- 176
QY 150 EANKALKFDEKAYGKRLDI-PSGNTL-----RIGAGQTRKVKQ-LIPLGGSKKV----IGM 199
DB 177 -----AKVKRIEYKGFVSEFLFGKQALHLHSIQISQERINKVEDVLKIGDTIEVKITEIDK 231
QY 200 NGLVNNIADERHKHAKDKAKS 221
DB 232 QGRVN-----ASHKVLEQSKN 247

RESULT 4
US-10-510-386-108
; Sequence 108, Application US/10510386
; Publication No. US20050244922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Olsen, Peter Bjarke
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294.204-US
; CURRENT APPLICATION NUMBER: US/10/510,386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 108
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-510-386-108

Query Match      6.6%; Score 77.5; DB 1; Length 259;
Best Local Similarity 21.7%; Pred. No. 0.71;
Matches 38; Conservative 36; Mismatches 54; Indels 47; Gaps 9;

QY 46 EARRGKKTVAQLMEECHMFLKDKDEVMPGVGMVDPDLGVBEATFPDGTGLVTNVNPIEDPH 105
DB 95 QKEKDKQSLKKIQEQVNRFFIKENLQKQVNTKLTDEGLLS-----IEDNIF 141
QY 106 FXAGEVKFGCKDKDIELNAGKEVTELEVNTGPKSLHVGSHF-----HFFEANKAL--- 155
DB 142 FDSGKAIR-QQDIPL--AKEVSDLLVLP-PRNIVISGHTDNVPINRSQPKSNWHL SVM 197
QY 156 -----KFDRE-----KAYGKRLDIPSGNTLRIGAGQTRKVKQ--LIPLG 191
DB 198 RAVNFWGLLIENPKLDAKIFSAGYGEFKPIASNDT-EEGRKRRVEILLIPIG 251

RESULT 5
US-10-689-742-50
; Sequence 50, Application US/10689742
; Publication No. US20050250180A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M
; APPLICANT: Lavallie, Edward R
; APPLICANT: Racie, Lisa A
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 00766.000091.10
; CURRENT APPLICATION NUMBER: US/10/689,742
; CURRENT FILING DATE: 2003-10-22
; PRIOR APPLICATION NUMBER: 09/746,783
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 231
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 50
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-689-742-50

Query Match      6.5%; Score 76.5; DB 1; Length 428;
Best Local Similarity 21.7%; Pred. No. 1.8;
Matches 55; Conservative 28; Mismatches 82; Indels 89; Gaps 14;

QY 3 LTPKEQEKFLYYAGEVARKKAE--GLKLNQPEAIAIYSAHINDEARRGKKTVA----- 55
DB 186 LTP--NTPFLIDNDGNIENTEKLQYSGERLYKFTVTAYDC-----GKRAADDAEV 234
QY 56 --QLMEBC----MHFLKKDEVMPGVGMVDPDLGVBEATFPDGTGLVTNV---WPIEDPHF 106
```

Db 235 EIQVPTKPSQGNWNRLEYAPGAGSL-----ALFP-GIRLETCDPLNLIQATIEL 286
QY 107 KAGEVFGCDKX-----IELNAGKEVTELE-VTNEGPK-----SLHV---GSHFHF 149
Db 287 QTSYVAKGCDRDNYSERALKLGGATGEVDLLPMGPNANWTAGLSVHSQDSSLIYWF 346
QY 150 EANKALKFDEKAYKRLDIPSGNLTIRIGAGQTRKVLIPGSKVKVGMNGLVNNIAD 209
Db 347 NGTQAVQ-----VPLGGP-----SGLSGPQDS 369
QY 210 RHKKKALDKAKSHG 223
Db 370 LSDHFTLSFWMKHG 383

RESULT 6
US-10-821-234-1089
; Sequence 1089, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 1089
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1089

Query Match 6.2%; Score 73.5; DB 1; Length 449;
Best Local Similarity 21.2%; Pred. No. 3.8; Indels 61; Gaps 9;
Matches 51; Conservative 35; Mismatches 93
QY 1 VKLTPKEQKFLYYAGEVARKKAGKLNQPEAIAYISAHIMDEARRGKKTVAOLMEE 60
Db 142 VKETPOQKQRLHVEQELTTE-----VEKITTVESATSEKLTVPVLLAQ 198
QY 61 CMHFLKKDEMPGVGNMV-PDLGVEATFPDGTGLVTNNWPIBPDEHFKAQVKGCDKDI 119
Db 189 -LAALKQQLVASHLEKILGPDAAINLTDPDGALAKRLLQLLEATKNSKG- 237
QY 120 ELNAGKEVTELVNTEGPKSLVHSHFH-----PFEANKALKFDR---EKAYGKRLDI 169
Db 238 --SGGK-----TTGTTPPSSLYTYELHSRPEQDKFSQAQVAELSKRLTELETAVRC 289
QY 170 PSGNLTIRIGAGQTRKVLIPGSK-----KVIGMGNLVNNIADERHKHKA 215
Db 290 DAQNPLSAGQACLMETVELLOAKVSLDALDLAVLDQVEARLQSVLGKVNIEA----KHKA 345

RESULT 7
US-10-858-730-6
; Sequence 6, Application US/10858730
; Publication No. US20050255568A1
; GENERAL INFORMATION:
; APPLICANT: Bailey, Richard B.
; APPLICANT: Blomquist, Paul
; APPLICANT: Doten, Reed
; APPLICANT: Driggers, Edward M.
; APPLICANT: Madden, Kevin T.
; APPLICANT: O'Leary, Jessica
; APPLICANT: O'Toole, George
; APPLICANT: Trueheart, Joshua
; APPLICANT: Walbridge, Michael J.

; APPLICANT: Yorgey, Peter S.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR AMINO ACID
; FILE REFERENCE: 14184-030001
; CURRENT APPLICATION NUMBER: US/10/858,730
; CURRENT FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/475,000
; PRIOR FILING DATE: 2003-05-30
; PRIOR APPLICATION NUMBER: US 60/551,860
; PRIOR FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Shewanella oneidensis
US-10-858-730-6

Query Match 6.1%; Score 72; DB 1; Length 418;
Best Local Similarity 21.7%; Pred. No. 4.7;
Matches 50; Conservative 34; Mismatches 82; Indels 64; Gaps 12;
QY 27 GLKLNQPEAIAY-----ISAHIMDEARRGKKTVAOLMEECHMFLKDKDEVMPGVGNVPDL 81
Db 216 GAKVLHPDVSVEYAOQRFKVPURVLSSEFAGQGTLLQFGDESELAWAASVQGIANKALATL 275
QY 82 GVEATFPDGTGL-----VTNNW--PIEPDE-----HFKAQVKGCDKDI 121
Db 276 TIEGLFTSSERYQALLACLARLEVDVFITPLKLNELSPVESVFMLEAKV----DILL 331
QY 122 NAGKEVTELVNTEGPKSLVHSHFHFEANKALKFDEKA-----YGRKLDIPSG---NT 174
Db 332 H-----EVLSE---SLDLGQ-----LIVERQRAKVSILVKGLOAKVGLLTGM 372
QY 175 LRIGAGQTRKVLIPGSKVKVGMNGLVNNIADERHKHKAIDKAKSHGF 224
Db 373 LDVLGNETIHAKLSTSEK-----LSTVIDERDLHKAV-RALHHAF 413

RESULT 8
US-10-793-626-1248
; Sequence 1248, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 1248
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-1248

Query Match 6.0%; Score 71; DB 1; Length 311;
Best Local Similarity 18.9%; Pred. No. 3.9;
Matches 58; Conservative 35; Mismatches 100; Indels 114; Gaps 13;
QY 1 VKLTPKEQKFLYYAGEVARKKAGKLNQPEAI--AVISAHIMDEARR-----49
Db 6 LKNLPSEESYFLFY-----INRPSIIVGNKQNTIEVNQAYIDKKQIDVVRRLISGGGAVYH 61
QY 50 --GKKTVAOLMEECHMFLK--KDEVMFGVGNVPDLGVGEATFP-----DGTKL 93
Db 62 DTGNLNFSPITDDDGHSFHNFKFTMP-IVQALQSMGVNAEMTGRNDIQVGAKISGNAM 120

QY 94 VTVNWPIDPEHFKAGEVKKFGCD-----KDIELNAGK-----125
Db 121 VKVK-----NRMFSGHTLMLNCDLNEVOKALKVNPAKIKSGVKSVKRKVANIEBPLEQP 175
QY 126 -----EVTLEVTNPGKSLHVGSHFHP-----FEANKALKFDRK 161
Db 176 IDIEBKILKIITIGENEVEBYIILTEEDWNKIKQLSDKEYRTWENYGSNPKYNIEREE 235
QY 162 AYK-----RLDPSGNTLR-----ICAGQTRKVQLIPLGSKKVIWMGMLVNNIAD 209
Db 236 KEKGFQIKLDVKKGRIERAKLFGDFGEGDVTELE-----HALVGLCHDF 282
QY 210 RHKHAL 216
Db 283 EHIEEAL 289

RESULT 9
US-10-793-626-1460
; Sequence 1460, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1460
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-1460

Query Match 5.9%; Score 70; DB 1; Length 329;
Best Local Similarity 25.2%; Pred. No. 5.3;
Matches 52; Conservative 19; Mismatches 89; Indels 46; Gaps 10;
QY 42 HIMDEARRGKTTVAQLMEECMHFLKDEVMVPGVGNMVPDLGVEATFPDGT-----KLVTVN 97
Db 85 YIKDITRRGKVPi-----IAGGTGLYIQSLLYNAYAFEDESISGDKMKQVK 129
QY 98 WPIEPEH-----KAGEVKFGCD-----KDIELNAGKEVTE-LEVTNPGKSLHVGSHFHP 149
Db 130 LKLKELEHNNKLEHLYLASFKESAKDHPNNRKEVLAIEYLYLTKKLLSRKKVQOP 189
QY 150 EANK-----ALKPDRKAY-----GKRLDPSGNTLRIGAGQTRKVQ-LIPLG-----GSKKV 196
Db 190 TENYDTLLIGIEMSRETLRLINKRVDI-----MLGHGLENEVQHLVQGGPEASQMQA 243
QY 197 IGWNLVNNIADIERHKHAKLDIAKSH 222
Db 244 IGYKELVPVIGKNISMENAVEKLKH 269

RESULT 10
US-10-821-234-1608
; Sequence 1608, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A

; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 1608
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1608

Query Match 5.9%; Score 70; DB 1; Length 437;
Best Local Similarity 20.5%; Pred. No. 7.8;
Matches 46; Conservative 31; Mismatches 65; Indels 82; Gaps 11;
QY 4 TPKEQEKELLYAGEVARKKAEGLKLNQPEATAYISAHIMDEARRGK--KTVAQLMEEC 61
Db 46 TPEFLRKP---PAGKVPAPFEGDDGFCVFESNAIAY---YVSNEELRGSTPEAAQV--QW 98
QY 62 MHFLKKDEVMPGVGNMVPDLGV-----EAT 86
Db 99 VSEADSDIVPPASTWVPPTLIGIMHNNKQATENAKEEVRILGLLDAYLKTFTFLVGERVT 158
QY 87 FPDGTKLVTVNW----PIEP-----DEHFKA--GEVFGCKDKDIELN 122
Db 159 LADITVVTLLWLYKQVLEPSFQAFNTNRWELTCINQPFRAVLGEVKL-CEKMAQFD 217
QY 123 AGKEVTELEVTNPGKSLHVGSHFHFPEANKALKFDEKAYGKR 166
Db 218 A-KKFAETOPKDTPRK-----EKSREEKQKPOAER 248

RESULT 11
US-10-793-626-244
; Sequence 244, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 244
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-244

Query Match 5.8%; Score 69; DB 1; Length 501;
Best Local Similarity 20.0%; Pred. No. 12;
Matches 42; Conservative 38; Mismatches 74; Indels 56; Gaps 9;
QY 30 LNOPEAIAYISAHIMDEARRGKTTVAQL-----MEECMHFLKKDEVMPG 73
Db 304 VTEPDEIRDSLAIVLDESRLNRLVNNELNVARMDAGLSVEKELOPTQIHLDDKWKSKYR 363
QY 74 VGNMVPDLGVEATFP-----DGTKLVTNNWPIEPDEHFKAGE-VKFGCKDIE 120
Db 364 MQS--BELGLTMTFDSNNDEQLWNYDMRMDQVLTNLDNATRYTQAGDSIKISIDSD 421
QY 121 LNAGKEVTELEVTNPGKSLHVGSHF-HFPEANKALKFDR-----KAYGKR 166
Db 422 FN-----ILATIDTGTGTIAPHELKQVDFRKYVDAARKRGKQGTGLGLFICKWIIIEHGG 477
QY 167 LDIPSGNLTIRIGAGQTRKVQLIPLGSKKV 196

| | | | |
|----|-----|--|-----|
| Qy | 107 | KAGEVKGCDKDIELNAGKEVTE-----LEVNTNEGPKSLHVGSHFFHFEANKALKFDREK | 16 |
| Db | 628 | KDGTVTFTGLSQDSGLTICKSTLNNDGLTVKDTNE--QIQVGAN-----GIKF--TN | 674 |
| Qy | 162 | AYGKRLDIPSGNTLRI-----CAGOTRK-----VOLIPLG--GSKK | 195 |
| Db | 675 | VGNSNPGTGANTARIITRDKIGFAGSDGAVDTNKPYLDQDKLVGNVKITNTGINAGGKA | 734 |
| Qy | 196 | VIGMGLVNNIADRRHKHAL-----DKAKSH | 222 |
| Db | 735 | ITGLSPTLPSIADQSSRNIELGNTIQDKDKN | 766 |

RESULT 14
 US-11-013-759-4
 ; Sequence 4, Application US/11013759
 ; Publication No. US20050249747A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Loosmore, Sheena M.
 ; APPLICANT: Sasaki, Ken
 ; APPLICANT: Yang, Yan Ping
 ; APPLICANT: Klein, Michel H.
 ; TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE
 ; TITLE OF INVENTION: PROTEIN OF MORAXELLA
 ; FILE REFERENCE: 1038-921MIS:jb
 ; CURRENT APPLICATION NUMBER: US/11/013,759
 ; CURRENT FILING DATE: 2004-12-16
 ; PRIOR APPLICATION NUMBER: US/09/361,619
 ; PRIOR FILING DATE: 1999-07-27
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 4
 ; LENGTH: 2047
 ; TYPE: PRT
 ; ORGANISM: Moraxella catarrhalis
 US-11-013-759-4

| | | | | | |
|-----------------------|-----|--------------|---------------|------------|--------------|
| Query Match | | 5.8%; | Score 68.5; | DB 7; | Length 2047; |
| Best Local Similarity | | 25.0%; | Pred. No. 96; | | |
| Matches | 38; | Conservative | 18; | Mismatches | 47; |
| | | | | Indels | 49; |
| | | | | Gaps | 8 |

| | | | |
|----|-----|--|-----|
| Qy | 107 | KAGEVKGCDKDIELNAGKEVTE-----LEVNTNEGPKSLHVGSHFFHFEANKALKFDREK | 161 |
| Db | 683 | KDGTVTFTGLSQDSGLTICKSTLNNDGLTVKDTNE--QIQVGAN-----GIKF--TN | 729 |
| Qy | 162 | AYGKRLDIPSGNTLRI-----CAGOTRK-----VOLIPLG--GSKK | 195 |
| Db | 730 | VGNSNPGTGANTARIITRDKIGFAGSDGAVDTNKPYLDQDKLVGNVKITNTGINAGGKA | 789 |
| Qy | 196 | VIGMGLVNNIADRRHKHAL-----DKAKSH | 222 |
| Db | 790 | ITGLSPTLPSIADQSSRNIELGNTIQDKDKN | 821 |

RESULT 15
 US-11-013-759-7
 ; Sequence 7, Application US/11013759
 ; Publication No. US20050249747A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Loosmore, Sheena M.
 ; APPLICANT: Sasaki, Ken
 ; APPLICANT: Yang, Yan Ping
 ; APPLICANT: Klein, Michel H.
 ; TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE
 ; TITLE OF INVENTION: PROTEIN OF MORAXELLA
 ; FILE REFERENCE: 1038-921MIS:jb
 ; CURRENT APPLICATION NUMBER: US/11/013,759
 ; CURRENT FILING DATE: 2004-12-16
 ; PRIOR APPLICATION NUMBER: US/09/361,619
 ; PRIOR FILING DATE: 1999-07-27
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 7
 US-11-013-759-7

:|::||: :|:|:|:~::~
Db 478 IDVES-----ELGKGT\$---FIIRLPKSQI 500

RESULT 12
US-11-013-759-3
; Sequence 3, Application US/11013759
; Publication No. US20050249747A1
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Sasaki, Ken
; APPLICANT: Yang, Yan Ping
; TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE PROTEIN OF MORAXELLA CATARRHALIS
; FILE REFERENCE: 1038-92LMIS:jg
; CURRENT APPLICATION NUMBER: US/11/013,759
; PRIOR FILING DATE: 2004-12-16
; PRIORITY CLAIMING DATE: 1999-07-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1992
; TYPE: PRP
; ORGANISM: Moraxella catarrhalis
US-11-013-759-3

Query Match 5.8%; Score 68.5; DB 7; Length 1992;
Best Local Similarity 25.0%; Pred.No.92;
Matches 38; Conservative 18; Mismatches 47; Indels 49; Gaps 8;

QY 107 KAGEVFGCDKDIELNAGKEVT-----LEVNEGPKSLHVGSHFFFEANKALKFDREK 161
DB 628 KDGVTFGLSQDSGLTGIGSTLNNDGLTVKTNE----GIKF--TN 674

QY 162 AYGRDLIPSGNTLRI-----GAGQTRK-----VOLIFLG---GSKK 195
DB 675 VNGSNPGTGIANTARITRDKGIFAGSDGAVDTNKPYLDQDKLVGNVKITNTGINAGGA 734

QY 196 VIWMNLVNVIADERHKHAL-----DIAKH 222
DB 735 ITGLSPTLSIADQSRSNIETLGMTIQDKOKSN 766

RESULT 13
US-11-013-759-13
; Sequence 13, Application US/11013759
; Publication No. US20050249747A1
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Sasaki, Ken
; APPLICANT: Yang, Yan Ping
; TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE PROTEIN OF MORAXELLA CATARRHALIS
; FILE REFERENCE: 1038-92LMIS:jg
; CURRENT APPLICATION NUMBER: US/11/013,759
; PRIOR FILING DATE: 2004-12-16
; PRIORITY CLAIMING DATE: 1999-07-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 13
; LENGTH: 1992
; TYPE: PRP
; ORGANISM: Moraxella catarrhalis
US-11-013-759-13

Query Match 5.8%; Score 68.5; DB 7; Length 1992;
Best Local Similarity 25.0%; Pred.No.92;
Matches 38; Conservative 18; Mismatches 47; Indels 49; Gaps 8;

QY 107 KAGEVFGCDKDIELNAGKEVT-----LEVNEGPKSLHVGSHFFFEANKALKFDREK 161
DB 628 KDGVTFGLSQDSGLTGIGSTLNNDGLTVKTNE----GIKF--TN 674

QY 162 AYGRDLIPSGNTLRI-----GAGQTRK-----VOLIFLG---GSKK 195
DB 675 VNGSNPGTGIANTARITRDKGIFAGSDGAVDTNKPYLDQDKLVGNVKITNTGINAGGA 734

QY 196 VIWMNLVNVIADERHKHAL-----DIAKH 222
DB 735 ITGLSPTLSIADQSRSNIETLGMTIQDKOKSN 766

```
; LENGTH: 2047
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-11-013-759-7

Query Match      5.8%; Score 68.5; DB 7; Length 2047;
Best Local Similarity 25.0%; Pred. No. 96;
Matches 38; Conservative 18; Mismatches 47; Indels 49; Gaps 8;

QY 107 KAGEVKFGCDKIELNAGKEVTE-----LEVTNEGPKSLHVGSHPHFFFEANKALKFEDREK 161
Db   ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : |||
QY 683 KDGTVTFGLSQDSGLTIGKSTLNNDGLTVKDNE---QIQVGAN-----GIKF--TN 729
Db   ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : |||
QY 162 AYGRKLDIPSGNTRLRI-----GAGQTRK-----VQLPLG---GSKK 195
Db   ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : |||
QY 730 VNGSNPGTGIANTRITRDKIGFAGSDGAVDTNKPYLDDQDKLQVGNVKITNTGINAGGKA 789
Db   ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : |||
QY 196 VIGMGLVNNIADERHKHAL-----DKAKSH 222
Db   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 790 ITGLSPTLPSIADQSSRNIELGNTIQDKDSN 821
Db   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
```

Search completed: November 28, 2005, 08:25:57
Job time : 3.84635 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 28, 2005, 08:13:38 ; Search time 18.5013 Seconds
(without alignments)
1009.915 Million cell updates/sec

Title: US-09-904-994B-2

Perfect score: 1181

Sequence: 1 VKLTPKEQKFLLYAGEVA.....ADERHKHAKLDKAKSHGFIK 226

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/aaa/5 COMB.pep.*
- 2: /cgn2_6/prodata/1/aaa/6 COMB.pep.*
- 3: /cgn2_6/prodata/1/aaa/H COMB.pep.*
- 4: /cgn2_6/prodata/1/aaa/ECTUS_COMB.pep.*
- 5: /cgn2_6/prodata/1/aaa/RE_COMB.pep.*
- 6: /cgn2_6/prodata/1/aaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| 1 | 622 | 52.7 | 234 | 2 | US-09-338-920B-11 |
| 2 | 610.5 | 51.7 | 238 | 1 | US-08-920-095-2 |
| 3 | 610.5 | 51.7 | 238 | 2 | US-09-338-920B-5 |
| 4 | 610.5 | 51.7 | 238 | 4 | PCT-US96-05800-2 |
| 5 | 610.5 | 51.7 | 245 | 2 | US-09-431-705-4 |
| 6 | 603.5 | 51.1 | 238 | 2 | US-09-338-920B-7 |
| 7 | 600 | 50.8 | 237 | 1 | US-08-467-822-22 |
| 8 | 600 | 50.8 | 237 | 2 | US-08-432-697-22 |
| 9 | 600 | 50.8 | 237 | 2 | US-08-466-248-22 |
| 10 | 573.5 | 48.6 | 228 | 2 | US-10-012-819-154 |
| 11 | 572.5 | 48.5 | 237 | 1 | US-08-467-822-20 |
| 12 | 572.5 | 48.5 | 237 | 2 | US-08-432-697-20 |
| 13 | 572.5 | 48.5 | 237 | 2 | US-08-466-248-20 |
| 14 | 566.5 | 48.0 | 237 | 2 | US-09-338-920B-9 |
| 15 | 544.5 | 46.1 | 213 | 2 | US-09-338-920B-13 |
| 16 | 431.5 | 36.5 | 840 | 1 | US-08-467-822-25 |
| 17 | 431.5 | 36.5 | 840 | 2 | US-08-432-697-25 |
| 18 | 431.5 | 36.5 | 840 | 2 | US-08-466-248-25 |
| 19 | 281.5 | 23.8 | 100 | 1 | US-08-967-513-3 |
| 20 | 281.5 | 23.8 | 100 | 1 | US-08-687-645B-3 |
| 21 | 281.5 | 23.8 | 103 | 2 | US-09-489-039A-9261 |
| 22 | 269.5 | 22.6 | 99 | 1 | US-07-732-242C-1 |
| 23 | 266.5 | 22.6 | 107 | 2 | US-09-543-681A-6101 |
| 24 | 264.5 | 22.4 | 132 | 2 | US-09-134-001C-5049 |
| 25 | 261.5 | 22.1 | 100 | 1 | US-08-467-822-23 |
| 26 | 261.5 | 22.1 | 100 | 2 | US-08-432-697-23 |
| 27 | 261.5 | 22.1 | 100 | 2 | US-08-466-248-23 |

| | | | | | | |
|----|-------|------|-----|---|----------------------|-------------------|
| 28 | 261 | 22.1 | 145 | 2 | US-09-134-001C-4982 | Sequence 4982, Ap |
| 29 | 255.5 | 21.6 | 103 | 2 | US-09-328-352-5887 | Sequence 5887, Ap |
| 30 | 250.5 | 21.2 | 100 | 2 | US-09-602-777A-18 | Sequence 18, Appl |
| 31 | 250.5 | 21.2 | 125 | 2 | US-09-252-991A-26884 | Sequence 26884, A |
| 32 | 247 | 20.9 | 111 | 2 | US-09-710-279-116 | Sequence 116, App |
| 33 | 225.5 | 19.1 | 153 | 2 | US-09-543-681A-6085 | Sequence 6085, Ap |
| 34 | 224 | 19.0 | 106 | 1 | US-07-732-242C-2 | Sequence 2, Appli |
| 35 | 224 | 19.0 | 119 | 2 | US-09-328-352-5915 | Sequence 5915, Ap |
| 36 | 212.5 | 18.0 | 162 | 2 | US-09-602-777A-10 | Sequence 10, Appl |
| 37 | 209 | 17.7 | 106 | 1 | US-08-967-513-4 | Sequence 4, Appli |
| 38 | 209 | 17.7 | 106 | 1 | US-08-687-645B-4 | Sequence 4, Appli |
| 39 | 208 | 17.6 | 109 | 2 | US-09-489-039A-9266 | Sequence 9266, Ap |
| 40 | 206 | 17.4 | 109 | 1 | US-08-467-822-24 | Sequence 24, Appl |
| 41 | 206 | 17.4 | 109 | 2 | US-08-432-697-24 | Sequence 24, Appl |
| 42 | 206 | 17.4 | 109 | 2 | US-08-466-248-24 | Sequence 24, Appl |
| 43 | 185 | 15.7 | 137 | 2 | US-09-252-991A-26886 | Sequence 26886, A |
| 44 | 92 | 7.8 | 289 | 2 | US-09-134-001C-4415 | Sequence 4415, Ap |
| 45 | 90.5 | 7.7 | 790 | 1 | US-08-363-560-2 | Sequence 2, Appli |

ALIGNMENTS

RESULT 1

US-09-338-920B-11
; Sequence 11, Application US/09338920B
; Patent No. 6709851
; GENERAL INFORMATION:
; APPLICANT: Soman, Gopalan
; APPLICANT: Thomas, William D., Jr.
; APPLICANT: Monath, Thomas P.
; TITLE OF INVENTION: Stabilization of Helicobacter Urease
; FILE REFERENCE: 06132/023002
; CURRENT APPLICATION NUMBER: US/09/338,920B
; CURRENT FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: US 08/928,081
; PRIOR FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Helicobacter heilmannii
US-09-338-920B-11

| | | | | |
|-----------------------|------------------|--|-----------|-------------|
| Query Match | 52.7% | Score 622; | DB 2; | Length 234; |
| Best Local Similarity | 54.7%; | Pred. No. 2.9e-62; | | |
| Matches 123; | Conservative 41; | Mismatches 59; | Indels 2; | Gaps 2; |
| Qy | 1 | VKLTPKEQKFLLYAGEVARKKAEGLKLNQPEAIAVISAHIMDEARRGKKTVAQLMEE | 60 | |
| Db | 1 | MKUTPKELDKMLHYAGELAKQKAKGKILNYTBAVALISAHVWEAARAGKKSVDLQWE | 60 | |
| Qy | 61 | CMHFLKKDEYMPGVGNMVPDLGVEATFPDGTGLVTVMPIEP-DEHFHAGEYKFGCDKI | 119 | |
| Db | 61 | GRILLKADDVMPGVAHMHEVGIAGPDTGLVTIHTFVEAGSKLAPGEVILK-NEDI | 119 | |
| Qy | 120 | ELNAGKEVTELVNTEGKSLHVGSHFFFEANKALPDRKAYGKRLDIPSGNTLRGA | 179 | |
| Db | 120 | TLNAGKHAVLKVNKNGDRPVQVSGSHFFFEVFNKLLDPDRKAYGKRLDIASGTAVREP | 179 | |
| Qy | 180 | GQTRKVLQILPLGSKKVIYGNGLVNNIADERHKHAKLDKAKSHGF | 224 | |
| Db | 180 | GEEKTVELIDGNGKRIYGFNALVDROADHDKKALKRAKEKH | 224 | |

RESULT 2

US-08-920-095-2
; Sequence 2, Application US/08920095
; Patent No. 5837240
; GENERAL INFORMATION:
; APPLICANT: Cynthia K. Lee et al.
; TITLE OF INVENTION: MULTIMERIC, RECOMBINANT UREASE VACCINE

```
;
;
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,095
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/431,041
; FILING DATE: 28-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06132/013001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-920-095-2

Query Match      51.7%; Score 610.5; DB 1; Length 238;
Best Local Similarity 52.7%; Pred. No. 6.1e-61;
Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;

QY 1 VKLTPEQEKFLYYAGAEVARKKAEGLKLNQPEATAYISAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLTPEKDLMLHYAGELAKKKEGKIKLNYVEAVALISAHIMEERAGKKTAAELMQE 60
QY 61 CMHFLKDEYMPGVGNMVPDLGVEATFPDGTCLVTNNWPIEDPHFKAGEVKGCDKDIE 120
Db 61 GRTLLKPDVMDGVASMIHEVGIEAMFPDGTCLVTVTHTPIEANGKLVPGEL-FLKNEDIT 119
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFFPEANKALKFDEKAYGKRLDIPSGNTLRIAG 180
Db 120 INEGKAVSVKVNVDGDPVQIGSHFFHFFEVNRCCLDFDREKTFGKRLDIASGTAVRFEPG 179
QY 181 QTRKVLQPLGGSKKVGNGLVNNTADRRHKHKKALDKAKSHGF 224
Db 180 EKSVELIDIGNRRIFGFNALVDROADNESKIALHRAKERGF 223

RESULT 3
PCT-US96-05800-2
; Sequence 5, Application US/09338920B
; Patent No. 6709851
; GENERAL INFORMATION:
; APPLICANT: Soman, Gopalan
; APPLICANT: Thomas, William D., Jr.
; APPLICANT: Monath, Thomas P.
; TITLE OF INVENTION: Stabilization of Helicobacter Urease
; FILE REFERENCE: 06132/023002
; CURRENT APPLICATION NUMBER: US/09/338,920B
; CURRENT FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: US 08/928,081
; PRIOR FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5

Query Match      51.7%; Score 610.5; DB 4; Length 238;
Best Local Similarity 52.7%; Pred. No. 6.1e-61;
Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;

QY 1 VKLTPEQEKFLYYAGAEVARKKAEGLKLNQPEATAYISAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLTPEKDLMLHYAGELAKKKEGKIKLNYVEAVALISAHIMEERAGKKTAAELMQE 60
QY 61 CMHFLKDEYMPGVGNMVPDLGVEATFPDGTCLVTNNWPIEDPHFKAGEVKGCDKDIE 120
Db 61 GRTLLKPDVMDGVASMIHEVGIEAMFPDGTCLVTVTHTPIEANGKLVPGEL-FLKNEDIT 119
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFFPEANKALKFDEKAYGKRLDIPSGNTLRIAG 180
Db 120 INEGKAVSVKVNVDGDPVQIGSHFFHFFEVNRCCLDFDREKTFGKRLDIASGTAVRFEPG 179
QY 181 QTRKVLQPLGGSKKVGNGLVNNTADRRHKHKKALDKAKSHGF 224
Db 180 EKSVELIDIGNRRIFGFNALVDROADNESKIALHRAKERGF 223

RESULT 4
PCT-US96-05800-2
; Sequence 2, Application PC/TUS9605800
; GENERAL INFORMATION:
; APPLICANT: Oravax, Inc.
; TITLE OF INVENTION: MULTIMERIC, RECOMBINANT UREASE VACCINE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/05800
; FILING DATE: 23-APR-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/431,041
; FILING DATE: 28-APR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/568,122
; FILING DATE: 06-DEC-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06132/020001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
PCT-US96-05800-2
```



```
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 237 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-467-822-22

Query Match      50.8%; Score 600; DB 1; Length 237;
Best Local Similarity 52.7%; Pred. No. 9.4e-60;
Matches 118; Conservative 43; Mismatches 61; Indels 2; Gaps 2;

QY 1 VKLTPEQEKFLLYAGGEVARKKAEGLKLNQPEALAIYSAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLTPKELDK-LMHYAGELAKRKGIGKLNVEAVALLSAHIMEERAGKKTAAELMQE 59
QY 61 CMHFLKDEVMGVMGNNVDPDLGVEATFPDGTKLVTNNWPIEPDHFHFKAGEVFGCDKDIE 120
Db 60 GRTLLKPDVDMGVASMIHEVGIEAMPDGTKLVTHTPIEANGKLVPGEL-FLKNEDIT 118
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFEANKALFKDREKAYGKRLDIPSGNTLRIGAG 180
Db 119 INEGKAVSVKVNVDPRVQIGSHFFHFEVNRCLDFDREKTFGKRLDIASGTAVRFEFG 178
QY 181 QTRKVLQIPLGSGKKVIGMGLVNNIADERHKKHAKDKAKSHGF 224
Db 179 EEKVELIDIGGNRRIFGFNALVDROADNESKKIALHRAKERGF 222

RESULT 8
US-08-432-697-22
; Sequence 22, Application US/08432697
; Patent No. 6248330
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-00000
```

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 237 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-432-697-22

Query Match      50.8%; Score 600; DB 2; Length 237;
Best Local Similarity 52.7%; Pred. No. 9.4e-60;
Matches 118; Conservative 43; Mismatches 61; Indels 2; Gaps 2;

QY 1 VKLTPEQEKFLLYAGGEVARKKAEGLKLNQPEALAIYSAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLTPKELDK-LMHYAGELAKRKGIGKLNVEAVALLSAHIMEERAGKKTAAELMQE 59
QY 61 CMHFLKDEVMGVMGNNVDPDLGVEATFPDGTKLVTNNWPIEPDHFHFKAGEVFGCDKDIE 120
Db 60 GRTLLKPDVDMGVASMIHEVGIEAMPDGTKLVTHTPIEANGKLVPGEL-FLKNEDIT 118
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFEANKALFKDREKAYGKRLDIPSGNTLRIGAG 180
Db 119 INEGKAVSVKVNVDPRVQIGSHFFHFEVNRCLDFDREKTFGKRLDIASGTAVRFEFG 178
QY 181 QTRKVLQIPLGSGKKVIGMGLVNNIADERHKKHAKDKAKSHGF 224
Db 179 EEKVELIDIGGNRRIFGFNALVDROADNESKKIALHRAKERGF 222

RESULT 9
US-08-466-248-22
; Sequence 22, Application US/08466248
; Patent No. 6258359
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,248
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/447,177
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
```

```

; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 237 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-466-248-22

Query Match 50.8%; Score 600; DB 2; Length 237;
Best Local Similarity 52.7%; Pred. No. 9.4e-60;
Matches 118; Conservative 43; Mismatches 61; Indels 2; Gaps 2;

QY 1 VKLTPEQEKFLYYAGEVARKKAEGLKLNQPEATAYISAHIMDEARAGKKTVAQLMEE 60
DB 1 MKLTPEKLDK-LMHYAGELAKKKEKGIKLNYYEAVALLISAHIMDEARAGKKTAAELMQE 59

QY 61 CMHFLKDEVMPCGVGNMVDPLGVEATFPDCTKLVTYNNWPIEPDEHFHFKAGVSKFGCDKIE 120
DB 60 GRTLLPKDDVMDGVASIMHEVGEATFPDGTGKLVTVHTPIEANGKLVPGEL-FLKNEDIT 118

QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDREKAYGKRLDIPSGNTRLIGAG 180
DB 119 INEGKKAHSVKKVKNVGDPRPQIGSHFHFPEVNRCLDFDREKTFGKRLDIASGTAVRFEG 178

QY 181 QTRKVLPIPLGSKKVIWMGLVNNIADERHKHKALDKAKSHGF 224
DB 179 EEKSVELIDIGNNRRIFGFNALVDROADNESKIALHRAKERGF 222

RESULT 10
US-10-012-819-154
; Sequence 154, Application US/10012819
; Patent No. 6916615
; GENERAL INFORMATION:
; APPLICANT: Legrain, Pierre
; APPLICANT: Selig, Luc
; APPLICANT: Rain, Jean-Christophe
; TITLE OF INVENTION: Collection of Prokaryotic DNA for Two-Hybrid System, Helicobacter
; TITLE OF INVENTION: pylori Protein-Protein Interactions and Applications thereof
; FILE REFERENCE: B5053
; CURRENT APPLICATION NUMBER: US/10/012,819
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: EP 99401066.8
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 278
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 154
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Helicobacter pylori
; US-10-012-819-154

```

[illegible]

```

Db      121  VKNVGRPVQIGSHFHFVFNRCILDFDREKTFGKRLDASGTAVRPFEPGEEKSVELIDIG 180
Qy      192  GSKKVGIMGNLVNNIADERHKHKLDAKSHGF 224
      181  GNRRIFGNALVDROADNESKKIALHRAKERGF 213

RESULT 11
US-08-467-822-20
; Sequence 20, Application US/08467822
; Patent No. 5843460
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITILE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; TITILE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,822
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 237 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..237
; OTHER INFORMATION: /note= "URE A - FIGURE 3."
US-08-467-822-20

```

```

Query Match      48.5%; Score 572.5; DB 1; Length 237;
Best Local Similarity 50.4%; Pred. No. 1.3e-56;
Matches 113; Conservative 43; Mismatches 67; Indels 1; Gaps 1;

QY 1 VKLTPEQKFLYYAGEVARRKKAELGLKNLOPEALAYISAHIMDSARRGKKTVAQLMEE 60
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 MKLTPEKDKIMWYHAGRLAEERLARGVKLNTVEALISGRVWEKARDGNKSVADLMQOE 60
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


[illegible]

```

RESULT 14
US-09-338-920B-9
; Sequence 9, Application US/09338920B
; Patent No. 6709851
; GENERAL INFORMATION:
; APPLICANT: Soman, Gopalan
; APPLICANT: Thomas, William D., Jr.
; APPLICANT: Monath, Thomas P.,
; TITLE OF INVENTION: Stabilization of Helicobacter Urease
; FILE REFERENCE: 06132/023002
; CURRENT APPLICATION NUMBER: US/09/338,920B
; CURRENT FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: US 08/928,081
; PRIOR FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-338-920B-9

```

| Query Match | 48.0% | Score 566.5 | DB 2 | Length 237 |
|-----------------------|-----------------|--|----------|------------|
| Best Local Similarity | 50.0% | Pred. No. 6.1e-56 | | |
| Matches 112 | Conservative 43 | Mismatches 68 | Indels 1 | Gaps 1 |
| QY | 1 | VKLTPEQEKFLLYAGAEVARKEKAEGLKLNQPEAIAYISAHIMDEARRGKTKTVAQILMEE | 60 | |
| DB | 1 | MKLTPEKLDKMLHYAGLAEALARGVKLNYTEAVALISGRVMEKARDNKSVDLMQE | 60 | |
| QY | 61 | CMHFLLKDEVMVGNNVPDGLVEATPDGTKLVTVNWPIDPEHFKAGEVKGCGKDIE | 120 | |
| DB | 61 | GRTWLKKNVMDGVSAMIEHVGTEANFPDGTGLVTTHTPTVEDNGKLAPGSV-FLKNEDIT | 119 | |
| QY | 121 | LNAGKEVTELEVTVNEGPKSLHVGSGHSHFFFEANKALKFDREKAYGCKRLDIPSGNTLRIGAG | 180 | |
| DB | 120 | INAGKEAISLUVKNKGDRPVQVQSGHSHFFFEVNKLLDDEAKSKFCRLDIASGTAVREPG | 179 | |
| QY | 181 | QTRKVQLIPLGSGKKVIGMNGLVNNTIADERHHKHKALDKAKSHGF | 224 | |
| DB | 180 | EEKSVELIDIGGNKRYTGFNSLYVDROADGKGLGKRAKEKF | 223 | |

```

RESULT 15
US-09-338-920B-13
; Sequence 13, Application US/09338920B
; Patent No. 6709851
; GENERAL INFORMATION:
; APPLICANT: Sonan, Gopalan
; APPLICANT: Thomas, William D., Jr.
; APPLICANT: Monath, Thomas P.
; TITLE OF INVENTION: Stabilization of Helicobacter Urease
; FILE REFERENCE: 06132/023002

```

```

; CURRENT APPLICATION NUMBER: US/09/338,920B
; CURRENT FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: US 08/928,081
; PRIOR FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 213
; TYPE: PRN
; ORGANISM: Helicobacter mustelae
US-09-338-920B-13

```

| | | | | | | |
|-----------------------|--------------|--|------------|-------|--------|------------|
| Query Match | 46.1%; | Score | 544.5; | DB 2; | Length | 213; |
| Best Local Similarity | 52.3%; | Pred. No. | 1.6e-53; | | | |
| Matches 112; | Conservative | 36; | Mismatches | 65; | Indels | 1; Gaps 1; |
| QY | 13 | LYVAGEVARKRKAEGKLQNQPAIAIYISAHIMDEARRGKKTVAQLMEECMHFLKDEVMP | 72 | | | |
| Db | 1 | LHYAGELAKRKXRGVKNLYVEAVALISWEIMEHAREGKKTVAIDLMSLGQVLKAEDVMD | 60 | | | |
| QY | 73 | GVCNMGVPLGVZATPPDGTKLVTWNVPITPDHFHFKAGEVKGCDKDIELNAGKEVTELEV | 132 | | | |
| Db | 61 | GVAALVHEYQVIEVCPDGGTLLVTVHNPIENNGKLHPGEFILK--DEDIVLNAGKEAIEVKV | 119 | | | |
| QY | 133 | TNEGPKSLHVGSHFFHFFFEANKALKFDREKAYGRKLDIPSGNTIRICAGOTRKVQLIPLGG | 192 | | | |
| Db | 120 | SNKGDRPIQVGHFFHFFETNKLLEFDREKRYGRRLDIASGTSVPFPGESKTVLLIQFGG | 179 | | | |
| QY | 193 | SKKVIQMGMLVNNIADERHKHKALDKAKSHGFIK | 226 | | | |
| Db | 180 | NORIFGNDLNNGOVNDNKRKALAAAKAGFIK | 213 | | | |

Search completed: November 28, 2005, 08:25:40
Job time : 19.5013 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 28, 2005, 08:00:02 ; Search time 36.4836 Seconds
(without alignments)
1497.963 Million cell updates/sec

Title: US-09-904-994B-3
Perfect score: 2999
Sequence: 1 MKMKQEVNTYGTGDKV.....KLCTSKPTSQVPLAQRYTFF 568

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|----------|--------------------|
| 1 | 2297 | 76.6 | 569 | 1 URKCBP | urease (EC 3.5.1.5 |
| 2 | 2244 | 74.8 | 569 | 1 S35291 | urease (EC 3.5.1.5 |
| 3 | 2028.5 | 67.6 | 569 | 2 C36950 | urease (EC 3.5.1.5 |
| 4 | 2016.5 | 67.2 | 571 | 2 F83681 | urease alpha subu |
| 5 | 1981.5 | 66.1 | 569 | 2 D69729 | urease (alpha subu |
| 6 | 1927.5 | 64.3 | 569 | 2 S75169 | urease (EC 3.5.1.5 |
| 7 | 1897 | 63.3 | 568 | 2 AG2264 | urease alpha chain |
| 8 | 1889.5 | 63.0 | 569 | 2 S47104 | urease (EC 3.5.1.5 |
| 9 | 1832 | 61.1 | 566 | 2 H83037 | urease alpha subu |
| 10 | 1826 | 61.1 | 568 | 2 C75586 | urease, alpha subu |
| 11 | 1826 | 60.9 | 572 | 2 H64075 | urease (EC 3.5.1.5 |
| 12 | 1825 | 60.9 | 567 | 1 C36138 | urease (EC 3.5.1.5 |
| 13 | 1822 | 60.8 | 570 | 2 S42607 | urease (EC 3.5.1.5 |
| 14 | 1813 | 60.5 | 567 | 1 D43719 | urease (EC 3.5.1.5 |
| 15 | 1809 | 60.3 | 567 | 1 S08480 | urease (EC 3.5.1.5 |
| 16 | 1809 | 60.3 | 568 | 2 G85654 | probable urease st |
| 17 | 1809 | 60.3 | 568 | 2 D90794 | urease alpha subu |
| 18 | 1799.5 | 60.0 | 838 | 2 A96699 | probable urease F1 |
| 19 | 1792.5 | 59.8 | 569 | 2 A97648 | urease alpha chain |
| 20 | 1792.5 | 59.8 | 569 | 2 AG2871 | urease alpha subu |
| 21 | 1785 | 59.5 | 570 | 2 AF3458 | urease (EC 3.5.1.5 |
| 22 | 1784.5 | 59.5 | 571 | 2 E90027 | urease alpha subu |
| 23 | 1780.5 | 59.4 | 571 | 2 S38485 | urease (EC 3.5.1.5 |
| 24 | 1779.5 | 59.3 | 572 | 2 AC0325 | urease (EC 3.5.1.5 |
| 25 | 1753.5 | 58.5 | 840 | 1 URJB | urease (EC 3.5.1.5 |
| 26 | 1715 | 57.2 | 573 | 2 A13332 | urease (EC 3.5.1.5 |
| 27 | 1710.5 | 57.0 | 835 | 2 T37939 | urease (EC 3.5.1.5 |
| 28 | 1708 | 57.0 | 573 | 2 S36028 | urease (EC 3.5.1.5 |
| 29 | 1696.5 | 56.6 | 598 | 2 G82890 | urease complex com |

RESULT 1

URKCBP

urease (EC 3.5.1.5) 62K chain - Helicobacter pylori (strains 26695, J99, and others)
N;Alternate names: urease beta chain; urease chain B; urease large subunit
C;Species: Helicobacter pylori
C;Date: 30-Sep-1991 #sequence revision 02-Dec-1994 #text change 05-Oct-2004
C;Accession: B38537; H64528; A71977; F41834; S07885; S12487; A49215; B61371; A35306; B41

R;Labigne, A.; Cussac, V.; Courcoux, P.
J. Bacteriol. 173, 1920-1931, 1991
A;Title: Shuttle cloning and nucleotide sequences of Helicobacter pylori genes responsib

A;Reference number: A38537; MUID:91161505; PMID:2001995
A;Accession: B38537
A;Molecule type: DNA

A;Residues: 1-569 <LAB>
A;Cross-references: UNIPROT:P14917; UNIPARC:UPI000002C20A; GR:M60398; NID:G149007; PIDN:
R;Tomb, J.P.; White, O.; Kervilange, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.

Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.

Nature 388, 539-547, 1997
A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A;Reference number: A64520; MUID:97394467; PMID:9252185
A;Accession: H64528
A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA
A;Residues: 1-569 <TOM>
A;Cross-references: UNIPARC:UPI000002C20A; GB:AE000529; GB:AE000511; NID:G2313152; PIDN:

A;Experimental source: strain 26695
R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;

Nature 397, 176-180, 1999
A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A;Reference number: A71800; MUID:99120557; PMID:9923682
A;Accession: A71977

A;Molecule type: DNA
A;Residues: 1-569 <ARN>
A;Cross-references: UNIPARC:UPI000002C20A; GB:AE001446; GB:AE001439; NID:G4154573; PIDN:

A;Experimental source: strain J99
R;Cussac, V.; Ferrero, R.L.; Labigne, A.
J. Bacteriol. 174, 2466-2473, 1992

A;Title: Expression of Helicobacter pylori urease genes in Escherichia coli grown under
A;Reference number: A41834; MUID:92210488; PMID:1313413
A;Accession: F41834

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 565-569 <CUS>

A;Cross-references: UNIPARC:UPI000002D2AD; GB:M84338
R;Clayton, C.L.; Pallen, M.J.; Kleanthous, H.; Wren, B.W.; Tabaqchali, S.
Nucleic Acids Res. 18, 362, 1990

A;Title: Nucleotide sequence of two genes from Helicobacter pylori encoding for urease
A;Reference number: S07884; MUID:90221820; PMID:2326167
A;Accession: S07885

| | | | |
|--|-----|---|-----|
| Db | 123 | AEGLIVTAGGIDTHIHIFISPOQIPTAFASGVTTWIMGGTCPADGTNATTITPGRANLKSM | 182 |
| Qy | 182 | LRAAEYSMNVLGLKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVAD EYD | 241 |
| Db | 183 | LRAAEYAMNLGLFLAKGNVSYBPSLRDQIEAGAIGFKIHEDWGSTPAAIHHCLNVADEYD | 242 |
| Qy | 242 | VQVCIHTDVTNEAGYVDDTLNANNGRAIHAYHIEGAGGSHSPDVTWAGELNLPSTTP | 301 |
| Db | 243 | VQVAIHDTLNEAGCVEDTL EATAGRTIHTFHTGAGGGHAPDVIKMAGESFNILPASTNP | 302 |
| Qy | 302 | TIPYTINTVAEHLDMLTCHHLDKRIREDLFOSQSRIRPGSIAAEDVLHDMGVAMTSSD | 361 |
| Db | 303 | TIPFTKNTAEHMDMLVMVCHHLDKS I KEDVQFADSRIRPOTIAAEQDLHDMGIFSI TSSD | 362 |
| Qy | 362 | SOAMGRAGEVIPRTWQTADKNKKEFGKLPEDGKDNDFRIKRYISKYITINPALTHGVSEY | 421 |
| Db | 363 | SOAMGRVGEVITRTWQTADKNKKEFGKLEKEGDNDFRIKRYISKYITINPGIAHGISDY | 422 |
| Qy | 422 | IGSVEEGKIADLVVNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPQPVYYREMPFGHH | 481 |
| Db | 423 | VGSVEVGKVALDLWSPAFGFKPKNIKGGFTALSQMGDANASIPTPQPVYYREMPFGHH | 482 |
| Qy | 482 | GKAKFDTSIITFVSKVAYENGVEKLGRLERQVLVKNCRNITKKPKFNDKTAKITVDPKT | 541 |
| Db | 483 | GKNKFDNTITFVSQAAYKAGIKBELGLDRAAPPVKNCRNITKKDLFPNDVTAHIDVNPET | 542 |
| Qy | 542 | FEVFDGKLCSTKPTSOVPLAQRYYTFP | 568 |
| Db | 543 | YKVKVDGKEVTSKADELSLAQLYNLF | 569 |
| RESULT 3 | | | |
| C36950 | | | |
| N:Alternate names: ureC protein | | | |
| C:Species: Bacillus sp. | | | |
| C:Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 05-Oct-2004 | | | |
| C:Accession: C36950 | | | |
| R:Maeda, M.; Hidaka, M.; Nakamura, A.; Masaki, H.; Uozumi, T. | | | |
| J. Bacteriol. 176, 432-442, 1994 | | | |
| A:Title: Cloning, sequencing, and expression of thermophilic Bacillus sp. strain TB-90 u | | | |
| A:Reference number: A36950; MUID:94117379; PMID:8288539 | | | |
| A:Accession: C36950 | | | |
| A:Status: preliminary | | | |
| A:Molecule type: DNA | | | |
| A:Residues: 1-569 <NAE> | | | |
| A:Cross-references: UNIPROT:Q07397; UNIPARC:UPI0000137D67; GB:D14339; NID:g393296; PIDN: | | | |
| C:Superfamily: urease, alpha subunit; urease 62K chain homology | | | |
| C:Keywords: hydrolase | | | |
| F:5-552/Domain: urease 62K chain homology <U62> | | | |
| Query Match 67.6%; Score 2028.5; DB 2; Length 569; | | | |
| Best Local Similarity 66.1%; Pred. No. 2.2e-137; | | | |
| Matches 375; Conservative 77; Mismatches 112; Indels 3; Gaps 2; | | | |
| Qy | 3 | MKKQEVNTYPTKGDKVLGDTDLWAEVEHDYTYTGEEELKFGAGKTIREGMGSN-SPD | 61 |
| Db | 5 | MRRQYADMFPGFVGDAIRLADSELFTIEIKDYTYTGDEVKFGGKVLRDGMGQHP LATS | 64 |
| Qy | 62 | ENTLDLVITNAMIIDYTYGIYKADIGIKNGKIHGIGKAGNKMDDGVSPHMVVGVTGEAL | 121 |
| Db | 65 | DECDVLVLTNAIIVDTYGIYKADIGIKDGMISIGKAGNPLMDGVD--NVIGAAETVIA | 122 |
| Qy | 122 | GEGMIITAGGIDSHTHFLSPQPPPTALANGVTTMFGGGTGPVDGTNATTITPKGNLHRM | 181 |
| Db | 123 | AEGMIVTAGGIDAHIFICPQIETALASGVTIMGGTGPATGTNATCTCPGFWNIHRM | 182 |
| Qy | 182 | LRAAEYSMNVLGLKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVAD EYD | 241 |
| Db | 183 | LQAAEEFFNLGLFLKGNCSDEAPLEQIEAGAVGLKLEHEDWGSTAAAI DTLKLVADRYD | 242 |
| Qy | 242 | VQVCIHTDVTNEAGYVDDTLNANNGRAIHAYHIEGAGGSHSPDVTWAGELNLPSTTP | 301 |

| | | | |
|---|-----|--|-----|
| Db | 243 | VQVAIHDTLNEGGFVEDTLKAIDGRVIHTYHTEGAGGGHAPDIIKAAGFPNLPSS TNP | 302 |
| Qy | 302 | TIPYTINTVAEHLDMLTCHHLDKRIREDLFOSQSRIRPGSIAAEDVLHDMGVAMTSSD | 361 |
| Db | 303 | TRPYTINTLEBHLDMVMVCHHL DANIPEDIAFADSRIRKETIAAEDVLHDLGVFSSMISSD | 362 |
| Qy | 362 | SOAMGRAGEVIPRTWQTADKNKKEFGKLPEDGKDNDFRIKRYISKYITINPALTHGVSEY | 421 |
| Db | 363 | SOAMGRVGEVITRTWQTADKNKQKGKQEDNGVDNFRVKRYIAKYITINPAIHAGIADY | 422 |
| Qy | 422 | IGSVEEGKIADLVVNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPQPVYYREMPFGHH | 481 |
| Db | 423 | VGSVEVEKGLADLVVNPAPFGVKDELVLKGMIAYSTMGDPNASIPTPQPVLYRMPFAAK | 482 |
| Qy | 482 | GKAKFDTSIITFVSKVAYENGVEKLGRLERQVLVKNCRNITKKPKFNDKTAKITVDPKT | 541 |
| Db | 483 | GDAKYQTSITFVSKAAVEKEGHEQLGKKVKPVHGIRKLTCKDILLNDKTPKIDVDPQT | 542 |
| Qy | 542 | FEVFDGKLCSTKPTSOVPLAQRYYTFP | 568 |
| Db | 543 | YEVKVDGQLVTCPEAETVPMAQRYFLF | 569 |
| RESULT 4 | | | |
| P83681 | | | |
| urease alpha subunit ureC [imported] - Bacillus halodurans (strain C-125) | | | |
| C:Species: Bacillus halodurans | | | |
| C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 05-Oct-2004 | | | |
| C:Accession: P83681 | | | |
| R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Hira | | | |
| Nucleic Acids Res. 28, 4317-4331, 2000 | | | |
| A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and | | | |
| A:Reference number: A83650; MUID:20512582; PMID:11058132 | | | |
| A:Accession: P83681 | | | |
| A:Status: preliminary | | | |
| A:Molecule type: DNA | | | |
| A:Residues: 1-571 <STO> | | | |
| A:Cross-references: UNIPARC:UPI00000C3801; GB:AP001507; GB:BA000004; NID:g10172612; PIDN: | | | |
| A:Experimental source: strain C-125 | | | |
| C:Genetics: | | | |
| A:Gene: ureC | | | |
| C:Superfamily: urease, alpha subunit; urease 62K chain homology | | | |
| Query Match 67.2%; Score 2016.5; DB 2; Length 571; | | | |
| Best Local Similarity 64.4%; Pred. No. 1.6e-136; | | | |
| Matches 368; Conservative 88; Mismatches 112; Indels 3; Gaps 2; | | | |
| Qy | 1 | MKKKQEVNTYPTKGDKVLGDTDLWAEVEHDYTYTGEEELKFGAGKTIREGMGS--N | 58 |
| Db | 1 | MKLTRAQHASLYGPTGDKVRLADTDLLELETKDYTYTGDECKGGKVLRDGMGQSAY | 60 |
| Qy | 59 | SPDENTLDLVITNAMIIDYTYGIYKADIGIKNGKIHGIGKAGNKMDDGVSPHMVVGVTGE | 118 |
| Db | 61 | TRDEGVLDLITNATIIDYTYGIYKADIGIKDGHIVIGIKGNPDIMDGVESHMVGASTE | 120 |
| Qy | 119 | ALAGEGMIITAGGIDSHTHFLSPQPPPTALANGVTTMFGGGTGPVDGTNATTITPKGNL | 178 |
| Db | 121 | AIAEGELIVTAGGIDAHIFISPOQIDVAIASGVTIMGGTGPATGTNATCTCTGKNWI | 180 |
| Qy | 179 | HRMLRAAEYSMNVLGLKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVAD | 238 |
| Db | 181 | ERMLEAADAPVNLGFLGKNASTPAPLREQIEAGAILGLHEDWGTTPAAIRTCLSVAD | 240 |
| Qy | 239 | EYDVQVCITDVTNEAGYVDDTLNANNGRAIHAYHIEGAGGSHSPDVTWAGELNLPSS | 298 |
| Db | 241 | RMDVQVAIHDTLNEAGFVEDTIKAIGDRVIHTYHTEGAGGGHAPDIMKVAGLPNVLES | 300 |
| Qy | 299 | TTPTIPTINTVAEHLDMLTCHHLDKRIREDLFOSQSRIRPGSIAAEDVLHDMGVAMT | 358 |
| Db | 301 | TNTPRPFTVNTIDSHLDMLVMVCHHLDPNVDPEDVAFADSRIRPETIAAEDIIQDLGVSMI | 360 |
| Qy | 359 | SSDSQAMGRAGEVIPRTWQTADKNKKEFGKLPED--GKDNDFRIKRYISKYITINPALTHG | 417 |

Db 361 SSDSQAMGRVGEVIIRTWQTADMKKQKQALAEQQKGNNDNRIRKRYVSKYITINPAITHG 420
Qy 418 VSEYIGSVBEGKIADLVVWNPFAFFGKPKIVIKGGMVVFSEMGDSNASVPTPQPVYVREM 477
Db 421 IDDYGVSVGKGLADLVWDPFRFGVKPELILKGLLIAYSQMGDPNASIPTPQPVSRPM 480
Qy 478 FQHHGKAKFDTSTITFVSKVAYENGVEKELGLERQVLVFNKCRNITKKDKFNKDKTAKITV 537
Db 481 FGSFGRARGTSTITFLSKAAMDGLGHEALGLQKIAHVNCRISIKRSKMKYNDATPNIEI 540
Qy 538 DPKEFEVFDGKLCCTSKPTSQVPLAQRYYTF 568
Db 541 DPTEYEVKVDGEMITCKPFEVALAQRYYFLP 571
RESULT 5
D69729
urease (alpha subunit) ureC - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 05-Oct-2004
C:Accession: D69729
R:Kunzt, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertea
C.; Bron, S.; Brouillet, S.; Bruchi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emmerson, P.F.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritzt, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A:Authors: Iaubert, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serd
akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosaato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, K.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumschein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:93844377
A:Accession: D69729
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-569 <KUN>
A:Cross-references: UNIPROT:P77837; UNIPARC:UPI000060B94; GB:Z99122; GB:AL009126; NID:G
A:Experimental source: strain 168
C:Genetics:
C:Gene: ureC
C:Superfamily: urease, alpha subunit; urease 62K chain homology
F:5-552/Domain: urease 62K chain homology <U62>
Query Match 66.1%; Score 1981.5; DB 2; Length 569;
Best Local Similarity 62.5%; Pred. No. 5.2e-134;
Matches 357; Conservative 95; Mismatches 114; Indels 5; Gaps 3;
Qy 1 MKMKQEVYNTYPTKDGKVRGLGDTDLWAEVHDYTYTYGEEKLFGAGKTIREGMQGSN-- 58
Db 1 MKMSREYAEFLPGTGTGDKIRLGDTDLWTEVEKDFTVYGEEMTFGGKTIKIDGMQGNRI 60
Qy 59 SPDENTLDLVITNAMIIDTYGIYKADIGIKNGKIHGIGKAGNKMDOGVSPhMVVGTEAL 118
Db 61 TKGQDGLDITVNNVLLDITGYIKADVGKDGIRVGVKSGNPDINDGVDPHVIVAGTE 120
Qy 119 ALAGEGMIITAGGIDSHTHFLSPQPPALANGVTMTFMGGTGPVDTGNATTITPPGKNWL 178
Db 121 VTSGEKILTAGGVDTTHIFICPQONEVALSSGVTTLLGGTGTPATGSKATTCTSGAWYM 180
Qy 179 HRMLRAAEYSVMNVGFLGKGNSSKKQLVEQVEBAGAIGFKLHEDWGTTPSAIDHCLSVAD 238
Db 181 ARMLEAAEEFPINVGFLGKNASDKAPLIEQVEAGAIGLKLHEDWGTTPSAINTCMEVVD 240
Qy 239 EYDVQVCIHTDVTNEAGYVDDTLNANNGRAIHAYHIEGAGGHSPPDVTWAGELNIPSS 298
Db 241 EADIQVAIHDTLINEAGFVETTTAALKNRVIHTYHTGAGGGHAPDIITKTCBELNVLPSSTN 300
Qy 299 TPTTPTINTVAEHLDMLTCHHLDKRREDIQLFSQSRIRPGSIAAEDVLHDMGVIAMT 358

Db 301 TPTTPTINTVAEHLDMLTCHHLDAKVPEDVAFSHSRAATAAEDILHDICALSMT 360
Qy 359 SSDSQAMGRAGEVIRPTWQTADKNKKEFGKLP-EDGKNDNFRIKRYISKYITINPALTHG 417
Db 361 SSDSQAMGRVGEVIIRTWQVADKMKQKQALAEQQKGNNDNRIRKRYIAKTIINPAITHG 418
Qy 418 VSEYIGSVBEGKIADLVVWNPFAFFGKPKIVIKGGMVVFSEMGDSNASVPTPQPVYVREM 477
Db 419 LSHEVGSVEKGLADLVWDPVFFGVKPELVKGMILARAQMGDPNASIPTPPEVFMBOQ 478
Qy 478 FGHGKAKFDTSTITFVSKVAYENGVEKELGLERQVLVFNKCRNITKKDKFNKDKTAKITV 537
Db 479 YASYGKANRSTSTITFMSQASIERGVAESLGEKRIKSPVKNIRKLSKLDKMLNSALPKIEI 538
Qy 538 DPKEFEVFDGKLCCTSKPTSQVPLAQRYYTF 568
Db 539 DPKYQVFADEELSCQPDYVPLGQRYFLP 569
RESULT 6
S75169
urease (EC 3.5.1.5) alpha chain - Synechocystis sp. (strain PCC 6803)
N:Alternate names: protein sll1750
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
o, K.; Okumura, S.; Shampo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S75169
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-569 <KAN>
A:Cross-references: UNIPROT:P73061; UNIPARC:UPI0000137D79; EMBL:D90903; GB:AB001339; NID
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
C:Gene: ureC
C:Superfamily: urease, alpha subunit; urease 62K chain homology
C:Keywords: hydrolase; metalloprotein; nickel
F:5-552/Domain: urease 62K chain homology <U62>
F:136,138,219,362/Binding site: nickel 2 (His, His, Lys, Asp) #status predicted
Query Match 64.3%; Score 1927.5; DB 2; Length 569;
Best Local Similarity 62.3%; Pred. No. 3.8e-130;
Matches 354; Conservative 97; Mismatches 114; Indels 3; Gaps 2;
Qy 2 MKKKQEVYNTYPTKDGKVRGLGDTDLWAEVHDYTYTYGEEKLFGAGKTIREGMQGSN-SP 60
Db 4 RMDRHSYAHTFGPTGTGDKVRGLADTELFIEQDYATYGVDEYKFGGKIVIRDMGQSPSLR 63
Qy 61 DENTLDLVITNAMIIDTYGIYKADIGIKNGKIHGIGKAGNKMDOGVSPhMVVGTEAL 120
Db 64 AEGAVDVVITNALLDWMGIVKADVGIKNGRIYAGKAGNPHIQDNVS--IIIGPSTBAI 121
Qy 121 AGEEMIITAGGIDSHTHFLSPQPPALANGVTMTFMGGTGPVDTGNATTITPPGKNLHR 180
Db 122 AGEEMIITAGGIDAHVHFICPQIETALASGVTTLVGGTGTPAAGTAKTCTPGANNHR 181
Qy 181 MLRAAEYSVMNVGFLGKGNSSKKQLVEQVEBAGAIGFKLHEDWGTTPSAIDHCLSVAD 240
Db 182 MLQAADFPINLGLGKNGSQPAALAEQIKAGAIGLKLHEDWGTTPAAIDNCLGVAEDY 241
Qy 241 DVQVCIHTDVTNEAGYVDDTLNANNGRAIHAYHIEGAGGHSPPDVTWAGELNIPSS 300
Db 242 DVQVAIHDTLINEAGFVETTTAALKNRVIHTYHTGAGGGHAPDIITKTCBELNVLPSSTN 301
Qy 301 TPTTPTINTVAEHLDMLTCHHLDKRREDIQLFSQSRIRPGSIAAEDVLHDMGVIAMTSS 360

Db 302 PTRPYTINTLEHLDMLMCHHLHRLNIPEDVAFASERIRRETTAAEDILHDLGAFSISS 361
QY 361 DSQAMGRAGEVIRPTWTQADKNKKEFGKLPEDGKNDNFRIKRYISKYTNINPALTHGVSE 420
Db 362 DSQAMGRVGEVICRTWTQAHKMKVQRGOLFPGETGNNDNLRAKYVAKYTNINPALTHGISE 421
QY 421 YIGSVBEGKIADLVVNPAPFVGKPKIVIKGMVVFSEMGDSNASVPTPOPVVYREMFHGH 480
Db 422 EIGSVBEGKIADLVVNPAPFVGKPKIVIKGMVVFSEMGDSNASVPTPOPVVYREMFHGH 481
QY 481 HGKAKFDTSTFTFSKVAYENGVEKGLGLERQVLPVKNCRNITKKDKFKNFKTAKITVDPK 540
Db 482 YGGAIAATSTFTFSKVAYENGVEKGLGLERQVLPVKNCRNITKKDKFKNFKTAKITVDPK 541
QY 541 TPEVFDGKLTCKTSKPTSOVPLAQRYTFF 568
Db 542 TYEVADGELLTCEPATVLPMAQRYFLF 569

RESULT 7
AG2264
urasee alpha chain [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 05-Oct-2004
C:Accession: AG2264
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; UID:21595285; PMID:11759840
A:Accession: AG2264
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-568 <KUR>
A:Cross-references: UNIPARC:UPI00000CE887; GB:BA000019; PIDN:BA875369.1; PID:g17132803;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr3670
C:Superfamily: urease, alpha subunit; urease 62K chain homology

Query Match 63.3%; Score 1897; DB 2; Length 568;
Best Local Similarity 61.4%; Pred. No. 5.8e-128;
Matches 349; Conservative 95; Mismatches 120; Indels 4; Gaps 3;

QY 2 KMKKQEVNTYGTGDKVRLGDTDLMAEVEHDYTTTGYBELKFGAGKTIREGMQSNPD 61
Db 4 RMSRQVAYETGPTVGRIRLADTELFQVEQDFTTYGDEVKFGGKVIKRGMSQSPIAN 63

QY 62 -ENTLDLVITNAMIIDYTGKADIGIKNGKIHGKAGNKQMDGVSPHVMVVGTEAL 120
Db 64 ADGAVDLVITNALILDMWGVKADIGIKNGKIHGKAGNPIYQDHYD--IIIGPGTEAL 121

QY 121 AGSGMILTAGIDSHTHFLSPQFPPTALANGVTTFMGGGTGPVDGTNATTITPGKWNLR 180
Db 122 AGSGMILTAGIDSHTHFLSPQFPPTALANGVTTFMGGGTGPVDGTNATTITPGKWNLR 181

QY 181 MLRAAEYSNMNVGFLGKGNSSKKQVLEQVAGAGIKGLHEDHGTTPSAIDHCLSVAD 240
Db 182 MLQAADAFPNLGLGKGNASQPGVLEQVIFAGAGIKGLHEDHGTTPATIDTCLTVADEY 241

QY 241 DVQVCTHTDVNAGYVDDTLNAMNGRAIHAYHIEGAGGSHSPDVTMAGELNLPSTT 300
Db 242 DVQVAIHTDVLNAGYVDDTLNAMNGRAIHAYHIEGAGGSHSPDVTMAGELNLPSTT 301

QY 301 PTTPYNTINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAVTSS 360
Db 302 PTTPYNTINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAVTSS 361

QY 361 DSQAMGRAGEVIRPTWTQADKNKKEFGKLPEDGKNDNFRIKRYISKYTNINPALTHGVSE 420
Db 362 DSQAMGRVGEVIIRPTWTQAHKMKVQRGOLFPGETGNNDNLRAKYVAKYTNINPALTHGIAQ 420

QY 421 YIGSVBEGKIADLVVNPAPFVGKPKIVIKGMVVFSEMGDSNASVPTPOPVVYREMFHGH 480
Db 421 YIGSVBEGKIADLVVNPAPFVGKPKIVIKGMVVFSEMGDSNASVPTPOPVVYREMFHGH 480
QY 481 HGKAKFDTSTFTFSKVAYENGVEKGLGLERQVLPVKNCRNITKKDKFKNFKTAKITVDPK 540
Db 481 FAGARNATSTFTFSQALEREIPQOLGURKSAVAVSGTRQLTKQDKMLNDALPHIEVDSE 540
QY 541 TFEVFDGKLTCKTSKPTSOVPLAQRYTFF 568
Db 541 SYEVADGELLTCEPATVLPMAQRYFLF 568

RESULT 8
S47104
urasee (EC 3.5.1.5) 62K chain - Bacillus pasteurii
N:Alternate names: ureC protein
C:Species: Bacillus pasteurii
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S47104
R:Moersdorf, G.; Weinmann, P.; Kaltwasser, H.
submitted to the EMBL Data Library, June 1994
A:Description: Nucleotide sequence of three genes on a urease encoding DNA-fragment from
A:Reference number: S47104
A:Accession: S47104
A:Molecule type: DNA
A:Residues: 1-569 <MOE>
A:Cross-references: UNIPROT:P41020; UNIPARC:UPI0000137D66; EMBL:X78411; NID:g498708; PID:
C:Genetics:
A:Gene: ureC
A:Start codon: GTG
C:Superfamily: urease, alpha subunit; urease 62K chain homology
C:Keywords: heterodcamer; hydrolase; metalloprotein; nickel
F:3-552/Domain: urease 62K chain homology <U62>
F:136,138,219,362/Binding site: nickel 2 (His, His, Lys, Asp) #status predicted

Query Match 63.0%; Score 1889.5; DB 2; Length 569;
Best Local Similarity 59.1%; Pred. No. 2e-127;
Matches 337; Conservative 112; Mismatches 118; Indels 3; Gaps 2;

QY 1 MMKKQEVNTYGTGDKVRLGDTDLMAEVEHDYTTTGYBELKFGAGKTIREGMQSN-- 58
Db 1 MKINRQVAYESYGTGDRVRLADTDL-GEVEKDYIYLGDEVNFGGKVLREGMGENTGY 59

QY 59 SPDENTLDLVITNAMIIDYTGKADIGIKNGKIHGKAGNKQMDGVSPHVMVVGTE 118
Db 60 TRTENVDLLLTNALILDYTGKADIGKVGKGYIVGKGGNPDIMDGVTPNNIVGTATE 119

QY 119 ALAGEGMIITAGGIDSHTHFLSPQFPPTALANGVTTFMGGGTGPVDGTNATTITPGKWNL 178
Db 120 VIAAEGKIVTAGGIDTHVHFNPQVDVALANGITTLFGGGTGAEGSKATTTVPGPWN 179

QY 179 HRMLRAAEYSNMNVGFLGKGNSSKKQVLEQVAGAGIKGLHEDHGTTPSAIDHCLSVAD 238
Db 180 EKMLKSTEGLPINVGILGKGGSSAPIMEQIDAGAAGLKIHDWDGATPASIDRSLTVD 239

QY 239 EYDVQVCTHTDVNAGYVDDTLNAMNGRAIHAYHIEGAGGSHSPDVTMAGELNLPSS 298
Db 240 EADVQVAIHTDVLNAGYVDDTLNAMNGRAIHAYHIEGAGGSHSPDVTMAGELNLPSS 299

QY 299 TTPTPYNTINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAVT 358
Db 300 TNPTPRPTVTNTIDSHLDMLMCHHLKQNIPEVAFADSRIRPETIAAEDILHDLGIISM 359

QY 359 SSSDQAMGRAGEVIRPTWTQADKNKKEFGKLPEDGKNDNFRIKRYISKYTNINPALTHGV 418
Db 360 STDALAMGRAGEVIRPTWTQADKNKKEFGKLPEDGKNDNFRIKRYISKYTNINPALAQCM 419

QY 419 SEYTGVBEGKIADLVVNPAPFVGKPKIVIKGMVVFSEMGDSNASVPTPOPVVYREMF 478
Db 420 AHEVGSIEEGKADLVLEWPEKFFGVKADRVIKGGIIAYAQIGDPSASITPTPOVPMGRMY 479

QY 479 GHGKAFEDTSITFVSKVAYENGVEKGLGLERQVLPVKNCRNITKKDKFKNFKTAKITVD 538

Db 480 GTVGLDIHDTNITFMSKSIQQGVPAKGLKRICTVKNCRNIGKDKMKNWVTTDIDIN 539

Qy 539 PKTFEVFDGKLGKTSKPTSOVLAQRYTFF 568

Db 540 PTYEVKVDGEVLTCEPVKELPMAQRYFLF 569

RESULT 9

H83037

urease, alpha subunit - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 06-Jan-2003

C:Accession: H83037

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Berman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen

A:Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: H83037

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-566 <STO>

A:Cross-references: UNIPARC:UPI00000C5E4A; GB:AE004900; GB:AE004091; NID:g9951134; PIDN:1

A:Experimental source: strain PA01

C:Genetics:

A:Gene: ureC; PA4868

C:Superfamily: urease, alpha subunit; urease 62K chain homology

Query Match 61.1%; Score 1832; DB 2; Length 566;

Best Local Similarity 59.2%; Pred. No. 2.6e-123;

Matches 336; Conservative 88; Mismatches 142; Indels 2; Gaps 1;

Qy 1 MKMKQEVNTYGTGKGVRLGDTDLWAEVHDYTYGELKFGAGKTIREGMSGNSP 60

Db 1 MKISQAYADMFGTGVDRVRLADTLTVEVERDFTYGEVYKFGGKVRIDMGQSOLG 60

Qy 61 DENTLDLVTNAMIIDYTGKADIGIKNGKIHGKAGKMDQGVSPHVVGVCTEAL 120

Db 61 AAQVDTVTNALILDHMGVAVKADVLGKGRQAIGKAGNPDIQPGV--NIAIGAGTEVI 118

Qy 121 AGEGLMIITAGGIDSHTHFLSPQFPALANGVTMTFGGCTGPDGTGNATTITPGKNLHR 180

Db 119 AGEGLMIITAGGIDTHLHFICPQIEALMSGVTMTLGGTGPAAGTATCTSGPHMAR 178

Qy 181 MLRAAEYSNMVGLGKGNSSKKQLVQVEAGAIGFKLHEDWGTTPSAIDHCLSVAD 240

Db 179 MLQAADAFPMNIGFTGKGNASLPPLQEQVLAGAIGLKLHEDWGSTPAIDNCLVAERH 238

Qy 241 DVQVCIHTDTVNEAGYVDDTLNMGRAIHAYHIEGAGGHSPPDVTMAGELNLPSSIT 300

Db 239 DIQVAIHTDTLINESGEVETTLGAFKRTIHTVHTGAGGGHAPDIIKACGFANVLPSSTN 298

Qy 301 PTIPYTIINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVAMTSS 360

Db 299 PTPPTFTNTIDHLDMLVCHLDLPAIAEDVAFASRIRETIAAEDILHDLGAFMIS 358

Qy 361 DSQAMGRAGEVTPRTWQADKNKKEFGKLPEDGKNDNFRIRKRIYISKYTINPALTHGVSE 420

Db 359 DSQAMGRVGEVITRTWQADKNKQRRGLDGDGARDNFRARIYIAKYTIINPAITHGISH 418

Qy 421 YTGSVEEGKIADLVNWPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPOPVVYREMF 480

Db 419 EYGSVEAGKWADLVLRPAPFGVKPSLIILKGAIAASLMGIDNGSITPTPOPVHYRPMFAS 478

Qy 481 HGKAKFDTSTTFVSKVAYENGVEKGLERQVLVKNCRNITKKDKFNDKTKAKITVDPK 540

Db 479 YAGSRHATSLTTFVQAFAAGVPPQQLGLRKAIGVSGCRGVQKTDLIHNGLYLTIEVDAQ 538

Qy 541 TFEVFDGKLGKTSKPTSOVLAQRYTFF 568

Db 539 NYQVRADGQLWCPEADVLPMAQRYFLF 566

RESULT 10

C75586

urease, alpha subunit - Deinococcus radiodurans (strain R1)

C:Species: deinococcus radiodurans

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 05-Oct-2004

C:Accession: C75586

R:White, O.; Eisen, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; Fraser, C.M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: C75586

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-568 <WHI>

A:Cross-references: UNIPARC:UPI00000C1687; GB:AE001863; GB:AE001825; NID:g6460670; PIDN:1

A:Experimental source: strain R1

C:Genetics:

A:Gene: DRA0318

A:Map position: 2

C:Superfamily: urease, alpha subunit; urease 62K chain homology

Query Match 61.1%; Score 1832; DB 2; Length 568;

Best Local Similarity 59.0%; Pred. No. 2.7e-123;

Matches 337; Conservative 100; Mismatches 128; Indels 6; Gaps 3;

Qy 1 MKMKQEVNTYGTGKGVRLGDTDLWAEVHDYTYGELKFGAGKTIREGMSGNS- 59

Db 1 MKVSQQAYADLYGPTGVDRVRLGDTTELLIEVERDLTTYGEVYKFGGKVRIDGLGQSSAA 60

Qy 60 --PDENTLDLVTNAMIIDYTGKADIGIKNGKIHGKAGKMDQGVSPHVVGVGT 117

Db 61 TRDANVPDLVITNALILDYTWGVKADVGKNGRISAIGKAGNPQTQGVTPGLTIAAST 120

Qy 118 EALAGEMIITAGGIDSHTHFLSPQFPPTALANGVTMTFGGCTGPDGTGNATTITPGKN 177

Db 121 EIVAGGLVLTAGVDTHIFIAFQQCWTALESGVTTMIGGCTGCTAGTSATCTPGQWH 180

Qy 178 LHRMLRAAEYSNMVGLGKGNSSKKQLVQVEAGAIGFKLHEDWGTTPSAIDHCLSVA 237

Db 181 IHRMLESAGLPLNFGLLGKGNASTQPPLAEQIRAGALGLKLHEDWGTTPAAIHAALSVA 240

Qy 238 DEYDVQVCIHTDTVNEAGYVDDTLNMGRAIHAYHIEGAGGHSPPDVTMAGELNLP 297

Db 241 EDYDVQVAIHTDTLINESGFVEDAIRAFAGRTIHTFTGAGGGHAPDIIRVAGLNPVLPS 300

Qy 298 STPTPTPTINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVAM 357

Db 301 STNPTPTPTINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVAM 360

Qy 358 TSSDSQAMGRAGEVTPRTWQADKNKKEFGKLPEDGKNDNFRIRKRIYISKYTINPALTHG 417

Db 361 MSSDSQAMGRVGEVITRTWQAQAHKMKVQGRPLAPDGR-ADNFRARYVAKYTINPAIAG 419

Qy 418 VSEYIGSVESGKIADLVNWPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPOPVVYRE 477

Db 420 ISHEVSGSVESGKIADLVNWPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPOPVVYRE 479

Qy 478 FGHGKAKFDTSTTFVSKVAYENGVEKGLERQVLVKNCRNITKKDKFNDKTKAKITV 537

Db 480 FAAYGCGPDATCLHFLVSQAGLEGGHLPDVG--RYSYAVKHTDRDIGKMDQLNAETPDIOV 537

Qy 538 DPKTFFVFDGKLGKTSKPTSOVLAQRYTFF 568

Db 538 NPETYEVRVNGELVTCFVDELPLAQRYFLF 568

RESULT 11

H64075

urease (EC 3.5.1.15) 62K chain - Haemophilus influenzae (strain Rd KW20)

N:Alternate names: urease alpha chain
C:Species: Haemophilus influenzae
C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
C:Accession: H64075
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.H.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, D.W.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
J.:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: H64075
A>Status: nucleic acid sequence not shown; translation not shown
Molecule type: DNA
Residues: 1-572 <TIGR>
A:CROSS-references: UNIPROT:P44391; UNIPARC:UIP0000137D6B; GB:U32736; GB:L42023; NID:g15
C:Superfamily: urease, alpha subunit; urease 62K chain homology
C:Keywords: hydrolase; metalloprotein; nickel
F:5-555/Domain: urease 62K chain homology <b62>
F:141,143,224,367/Binding site: nickel 2 (His, His, Lys, Asp) #status predicted

Query Match 60.9%; Score 1826; DB 2; Length 572;
Best Local Similarity 58.9%; Pred. No. 7.2e-123;
Matches 337; Conservative 93; Mismatches 136; Indels 6; Gaps 3;

Qy 1 MMKKQEYVNTYTPTGDKGVRLGDTDLWAEVEHDDTYTTTGEEELKFAGKTIREGMQSNS - 59
 :
Db 3 LTISRAQYVATYGTVTDVKVRLGDTNLWATEQLLTKGDECKFGGKSVRDMAQSGTA 62

Qy 60 -- PDENTLDLVTNAMIID-YTIYKADIGIKNGKIHGKAGNKMGGVSPhmVG 116
 :
Db 63 TRDNPNVLDFVTNMVIIDAALGLIIEKDGRIVGIGAGNPDTNDNVTPNNIIAS 122

Qy 117 TEALAGEMIITAGGDSHTHELSPQQFPFTALANGVTTFMGCGTGPVDGTWATTTPGW 176
 || :
Db 123 TVHNHAHLIAAGGDITHIFICPOQAQHIESGVTVLLIGGCTPADGTHATTCTPGAW 182

Qy 177 NLHRMLRAAEYSMNVFGLKGSSSKQVLEQVEAGAIGFKLHEDWGTTTPSAIDHCLSV 236
 :
Db 183 YMERMFQAAEALPVNVGFEGKGCNSTLDPLREQIEAGALGLKHEDWGATPAVIDSALKV 242

Qy 237 ADEYDVQCCHTDTVNEAGVVDDLNAWNGRAIHAYHIEGAGGHSPVITWAGELNILP 296
 :
Db 243 ADEMIDIQVAIHDTLTNESGLEDTMKAIIDGRVITHFTTEGAGGGHAPDIIRKAAMYSNLP 302

Qy 297 SSTTTIPTINTVAEHLDMLTCHLHKRIREDLPQSQRIRPSGIASEDLVHDMDGYIA 356
 :
Db 303 ASTNTRPPETKNITDSEHLMVMVCHLDKRVPEDEVAFADSIRPETIAAIEDILHDMGVFS 362

Qy 357 MTSSDSQAMGRAGEVIPRTWTADKNKEPKLPEDGKNDNFRIKYISKYITINPALTH 416
 :
Db 363 INSSDSQAMRGIEVVIRTQTADKNMQRGELNEG--NDNFRIKYIAKYITINPATAH 420

Qy 417 GVSEYIGSVGEKGIADLVVNMPAFFGVKPXIVIKGGMVVFSEMGSNASVPPTQPYYRE 476
 :
Db 421 GIAEHIGSLVGKIADIWLKWPFFGVKPEVWKPFISYAKMGDPNASIPTPQPFVFRP 480

Qy 477 MPHHGKAKFDTSITPVSKVAYENGVEKELGEROVLVKNCRNIKKDPFNENDTKAIT 536
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 481 MYGAQGGLATAQTAVFVSVQAAEKADRACFLGHKETIAVKGRCNRVGGKDLVHNDVTPNIT 540

Qy 537 VDPEKTFEPVDGKLCTSKPTSQVPLAQRYTF 568
 :
Db 541 VDAERYEVRDGELITCEPDVSDVPLGORYFLF 572

RESULT 12
C36138
urease [EC 3.5.1.5] 62K chain [validated] - Klebsiella pneumoniae
N:Alternate names: urea amidohydrolase; urease alpha chain; urease chain C
C:Species: Klebsiella pneumoniae
C>Date: 30-Nov-1990 #sequence_revision 02-Dec-1994 #text_change 05-Oct-2004
C:Accession: C36138

R.Mulrooney, S.B.; Hausinger, R.P.
J. Bacteriol. 172, 5837-5843, 1990
A>Title: Sequence of the Klebsiella aerogenes urease genes and evidence for accessory protein A
A:Reference number: A36138; MUID:91008957; PMID:2211515
A:Accession: C36138
A:Molecule type: DNA
A:Residues: 1-567 <MUI>
A:Cross-references: UNIPARC:UPI00001120B5; GB:M36068; NID:g149335; PIDN:AAA25151.1; PID:
submitted to the Brookhaven Protein Data Bank, June 1995
R:Jabri, E.; Karplus, P.A.
A:Reference number: A65977; PDB:JKRA
A:Contents: annotation; X-ray crystallography, 2.3 angstroms, residues 2-567
R:Jabri, E.; Carr, M.B.; Hausinger, R.P.; Karplus, P.A.
Science 268, 998-1004, 1995
A>Title: The crystal structure of urease from Klebsiella aerogenes.
A:Reference number: A56340; MUID:95273988; PMID:7754395
A:Contents: annotation; X-ray crystallography, 2.3 angstroms
R:Jabri, E.; Karplus, P.A.
Biochemistry 35, 10616-10626, 1996
A>Title: Structures of the Klebsiella aerogenes urease apoenzyme and two active-site mutants
A:Reference number: A58580; MUID:96346054; PMID:8718850
A:Contents: annotation; X-ray crystallography, 2.3 angstroms; active site mutagenesis
C:Genetics:
A:Gene: ureC
A:Complex: trimer of heterotrimeric subunits consisting of 11K (gamma), 12K (beta), and
C:Function:
A>Description: catalyzes the hydrolysis of one molecule of urea to carbon dioxide and two
C:Superfamily: urease, alpha subunit; urease 62K chain homology
C:Keywords: heterotrimer; hydrolase; metalloprotein; nickel
F:2-567/Product: urease 62K chain #status experimental <MAT>
F:4-550/Domains: urease 62K chain homology <U62>
F:134,136,217,360/Binding site: nickel 2 (His, Lys, Asp) #status experimental
F:217,246,272/Binding site: nickel 1 (Lys, His, His) #status experimental
F:217/Binding site: carbon dioxide (Lys) (covalent) (by urease activate) #status experim
F:219,320/Active site: His #status experimental

| | Query Match | Score | 1825; | DB 1; | Length | 567; | |
|----|-----------------------|---|--------------|--------|------------|------|----|
| | Best Local Similarity | 59.9%; | Pred. | No. | 8.4e-123; | | |
| | Matches | 339; | Conservative | 84; | Mismatches | 141; | |
| | | | | Indels | 2; | Gaps | 1; |
| Qy | 3 | MKKQEYVNTYGPTKGDKVRGLGDTDLWAEVEHDVTTYGEELKFGAGKTIREGMQSNDPE | 62 | | | | |
| Dd | 4 | ISRQAYADMFGPTGVGDVKRLADTELIEVEDDITTYGEERKVGKIVRDGMGGQOMLAA | 63 | | | | |
| Qy | 63 | NTLDLVITNAMIIDVTGIYKADIGIKNGKTHGIGKAGKDMQGVSPHMVVGVGTREALG | 122 | | | | |
| Dd | 64 | DCVDLVLTNALIDHWGIVKADIGVKDGRIFALGKAGNPDIQPNT--IPTGAATEVIAA | 121 | | | | |
| Qy | 123 | EGMIITAGGISHTFLSLPQQPFATALANGVTMTMGCGTGPVDGTNATITTPKKNLHRML | 182 | | | | |
| Dd | 122 | EKGIVTAGSIDTHIHWICPQAAREALVSGVTTWVGSGTGPAAGTHATTCTPGPWYSRML | 181 | | | | |
| Qy | 183 | RAAEYSNNVGFPLKGNSSKKQLVQVEAGATGFKLHEDWGTTPSAIDHCLSVADEYDV | 242 | | | | |
| Dd | 182 | QAADSLPNVIGLLKGNVSQPDALREQVAAGVIGLKITHEDWGATFAAIDCALTVADENDI | 241 | | | | |
| Qy | 243 | QVCILHTDTVNAGYVDDTLNANNGRAIHAYHI8GAGGHSPDVLTMAGELNILPSSTTPT | 302 | | | | |
| Dd | 242 | QUALHSDTLNESGFFVEDTLAALGGRTIHTFTBSGAGGHAPDIIITACNHPNLIPSSTNPT | 301 | | | | |
| Qy | 303 | IPTYINTVAEHLDMLTCHLLDKIRREDLQFSQRIRPGISAIEDVLHDMGVIAMTSSDS | 362 | | | | |
| Dd | 302 | LPTYLNTIDEHLDMVMCHLLDPDAEDVAFASRRIRRETIAAEDVLHDLGAFSLTSDDS | 361 | | | | |
| Qy | 363 | QAMGRAGNVIERTQTADKNKKEFGKL PEDGKONDNFRIKRYISKYTINPALTHGVSEYI | 422 | | | | |
| Dd | 362 | QAMGRVGEVILRTTWQVAHRMKVQRGALABETGDNDFRVKRYIAKYITNPALTHTGIAHEV | 421 | | | | |
| Qy | 423 | GSVEEGKTADEVVNPAEPFGVPKPIVIGKVVFSEMGDSNASVPTPOPYYRMEFGHHG | 482 | | | | |
| Dd | 422 | GSIEVGKLADLVVWSPPAFGVKPATPVIKGMGMAIAMPGDINAS IPTOPVHYRPMFWGALG | 481 | | | | |
| Qy | 483 | KAKEDPTSITFVSKVAYENGVEKGLIERQVL PVKNCRNITKKQPKFNKD KTAKITVDPKTF | 542 | | | | |

Db 492 SARHCRITFLSQAAANGVAERLNLRSIAIVVVGKRTVQKADMVHNSLQPNITVDAQTY 541
QY 543 EYFVDGKLTCKPTSOVPLAQRVTF 568
Db 542 EVRVDGELITSEPADVLPMAQRVFLF 567
RESULT 13
S42607
urease (EC 3.5.1.5) - Rhizobium melliloti
C:Species: Rhizobium melliloti
C:Date: 25-Dec-1994 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C:Accession: S42607
R:Mikech, G.; Arnold, W.; Lentzsch, P.; Priefer, U.B.; Puehler, A.
Mol. Gen. Genet. 242, 539-550, 1994
A:Title: A 4.6 kb DNA region of Rhizobium melliloti involved in determining urease and hy
ing frames.
A:Reference number: S42601; MUID:94166766; PMID:8121412
A:Accession: S42607
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-570 <NIK>
A:Cross-references: UNIPROT:P42885; UNIPARC:UPI000016FFCB; GB:S69145; NID:G545795; PIDN:
C:Genetics:
A:Gene: ureC
C:Superfamily: urease, alpha subunit; urease 62K chain homology
C:Keywords: hydrolase; metalloprotein; nickel
F:5-553/Domain: urease 62K chain homology <U62>
F:136,138,219,362/Binding site: nickel 2 (His, His, Lys, Asp) #status predicted
Query Match 60.8%; Score 1822; DB 2; Length 570;
Best Local Similarity 59.8%; Pred. No. 1.4e-122;
Matches 340; Conservative 95; Mismatches 130; Indels 4; Gaps 3;
QY 2 KMKKQBYVNTYPTKGDVRLGDTDLWAEVHDYTTYTGEBELFKGAGKTIREGQGQNSPD 61
Db 4 RMSRAAYANMFGPTGDKVRLADTELFIIEKDFTHGEVFKGGKVRDGMGQSVTR 63
QY 62 E-NTLDELVTNMIIDYTIYKADIGIKNGKIHGKAGNKMDDGVSPHVVGVGTEAL 120
Db 64 EGGAVDTVTNALILDHGVIKADIGLKGRIAAIGKAGNPDMPQGV--IIVGPGTEVI 121
QY 121 AGEGLMIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGGTGPVDGNTATTITPGKWNLR 180
Db 122 AGEGLVITAGGIDSHTHFLSPQFPPTALANGVTTMFGGGTGPVDGNTATTITPGKWNLR 181
QY 181 MLRAAEYSNMVGLFGKGNSSKKQLVEQVGAIGFKLHEDWGTTTPSAIDHCLSVAD 240
Db 182 MIEAADAPFNNLAFACKGNASLPALVEMVGLGATSLKLHEDWGTTTPAIDCCLSVADEY 241
QY 241 DVQVCIHDTVNEAGVYDDTLNAMNGRAIHAYHIEGAGGHSDDVITMAGELNLPSTT 300
Db 242 DVQVMIHTDNLNBSGFEVDIAAIGKRTIHAYHTEGAGGHHAPDIIIRICQPNVPSSTN 301
QY 301 PTIPYNTVAEHLDMMLTCHLDKRIREDLQFSQSRIKPGSIAAEDVLHDMGVIAWTS 360
Db 302 PTRPYNTVAEHLDMMLTCHLDKRIREDLQFSQSRIKPGSIAAEDVLHDMGVIAWTS 361
QY 361 DSQAMGRAGEVIRPTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYTTINPALTHGVSE 420
Db 362 DSQAMGRVGEVIRPTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYTTINPALTHGVSE 421
QY 421 YIGSVEEGKIADLVVWNPAPFFGVKPKIVIKGMVMVFSEMGDSNASVPTPQVYVREMF 480
Db 422 EIGSLEVGKRALVLNAPAFVGVKPKIVIKGMVMVFSEMGDSNASVPTPQVYVREMF 481
QY 481 HGKAKEDTSTITFVSKVAYENGVEKGLERQVLPVKNCR-NITKDFKFKNDKAKITVDP 539
Db 482 YGSRSTNSSTVTFVSPASLDAGRLGVAKELVAVQNTRGIGKASMIHNSLTPHIEVDP 541
QY 540 KTEVFVDGKLTCKPTSOVPLAQRVTF 568

Db 542 ETYEVRADGELLTCEPATVLPMAQRVFLF 570
RESULT 14
D43719
urease (EC 3.5.1.5) 62K chain - Proteus mirabilis
N:Alternate names: urease alpha chain
C:Species: Proteus mirabilis
C:Date: 03-Mar-1993 #sequence_revision 02-Dec-1994 #text_change 05-Oct-2004
C:Accession: D43719
R:Jones, B.D.; Mobley, H.L.T.
J. Bacteriol. 171, 6414-6422, 1989
A:Title: Proteus mirabilis urease: nucleotide sequence determination and comparison with
A:Reference number: A43719; MUID:90078080; PMID:2687233
A:Accession: D43719
A:Molecule type: DNA
A:Residues: 1-567 <JON>
A:Cross-references: UNIPROT:P17086; UNIPARC:UPI0000137D74; GB:M31834; NID:G150914; PIDN:
C:Genetics:
A:Gene: ureC
C:Complex: trimer of heterotrimeric subunits consisting of 11K (gamma), 12K (beta), and
C:Function:
A:Description: catalyzes the hydrolysis of one molecule of urea to carbon dioxide and two
C:Superfamily: urease, alpha subunit; urease 62K chain homology
C:Keywords: heterotrimer; hydrolase; metalloprotein; nickel
F:4-550/Domain: urease 62K chain homology <U62>
F:134,136,217,360/Binding site: nickel 2 (His, Lys, Asp) #status predicted
F:217,246,272/Binding site: nickel 1 (Lys, His, His) #status predicted
F:217/Binding site: carbon dioxide (Lys) (covalent) (by urease activase) #status predicted
F:219,320/Active site: His #status predicted
Query Match 60.5%; Score 1813; DB 1; Length 567;
Best Local Similarity 59.5%; Pred. No. 6.1e-122;
Matches 337; Conservative 91; Mismatches 136; Indels 2; Gaps 1;
QY 3 MKKQBYVNTYPTKGDVRLGDTDLWAEVHDYTTYTGEBELFKGAGKTIREGQGQNSPD 62
Db 4 ISQAVADMFGPTGDKVRLADTELFIIEKDFTHGEVFKGGKVRDGMGQSVSA 63
QY 63 NTLDELVTNMIIDYTIYKADIGIKNGKIHGKAGNKMDDGVSPHVVGVGTEALAG 122
Db 64 ECDVLITNMIIDYTIYKADIGIKNGKIHGKAGNPDMPQGV--IIVGPGTEVAG 121
QY 123 EGMIIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGGTGPVDGNTATTITPGKWNLRML 182
Db 122 EGKIVITAGGIDSHTHFLSPQFPPTALANGVTTMFGGGTGPVDGNTATTITPGKWNLRML 181
QY 183 RAAAEYSNMVGLFGKGNSSKKQLVEQVGAIGFKLHEDWGTTTPSAIDHCLSVAD 242
Db 182 EAVDELPIVNLFGKGCVSQPEAREQITAGAILGKIHEWDGATPMAIHNCINVADEMDV 241
QY 243 QVCIHTDTVNEAGVYDDTLNAMNGRAIHAYHIEGAGGHSDDVITMAGELNLPSTTPT 302
Db 242 QVAIHSDTLNEGGFYETVKAIGRVHVFTEGAGGHHAPDVIKSGEPNIPASTNPT 301
QY 303 IPTYTINTVAEHLDMMLTCHLDKRIREDLQFSQSRIKPGSIAAEDVLHDMGVIAWTS 362
Db 302 MPTYTINTVAEHLDMMLTCHLDKRIREDLQFSQSRIKPGSIAAEDVLHDMGVIAWTS 361
QY 363 QAMGRAGEVIRPTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYTTINPALTHGVSEYI 422
Db 362 QAMGRVGEVIRPTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYTTINPALTHGVSEYI 421
QY 423 GSVEEGKIADLVVWNPAPFFGVKPKIVIKGMVMVFSEMGDSNASVPTPQVYVREMF 482
Db 422 GSIEKGKRALVLNAPAFVGVKPKIVIKGMVMVFSEMGDSNASVPTPQVYVREMF 481
QY 483 KAKAKEDTSTITFVSKVAYENGVEKGLERQVLPVKNCR-NITKDFKFKNDKAKITVDP 542
Db 482 KAKYQTSMI FMSKAGIEAGVPEKGLGLSLIGRVEGCRHITKASMIHNNVPHIELDPQTY 541
QY 543 EVFVDGKLTCKPTSOVPLAQRVTF 568

Search completed: November 28, 2005, 08:20:42
Job time : 38.4836 secs

| Query Match | 60.3% | Score 1809; | DB 1; | Length 567; |
|---------------------------|--------|---|-----------|-------------|
| Best Local Similarity | 58.8%; | Pred. No. 1.2e-121; | | |
| Matches 333; Conservative | 94; | Mismatches 137; | Indels 2; | Gaps 1; |
| Qy | 3 | KKKQYVNTYGTGDKVRIGLDTDLMAEVEHDYTTYTGELKFGAGKTIREGMGQSNSPDE | 62 | |
| Db | 4 | ISRGAYADMEGPTTGDRLRLADTELFLIEQDFTTYGSEVKFGGKVRIDGMGQSQVSA | 63 | |
| Qy | 63 | NTLDLVITNAMIIDYTYGIKADYGIKNGKHIGICKAGNKMDDQVSPHVMVCGTEALAG | 122 | |
| Db | 64 | ECVDVLITNAIIIDHWIGIKADYGIKDGRIYGIKAGNPDPQPNVD--IVTGGPTEVAG | 121 | |
| Qy | 123 | EGMIITAGGIDSHTHFLSPQOFPPTALANGVTTFMGGGTGPVDGTNATITTPGKNLHRLM | 182 | |
| Db | 122 | ECKIITAGGVDTIHIFICPQAEGLISGVTTFIGGGTGPVAGTNATITTPGIMNHRML | 181 | |
| Qy | 183 | RAAEYSVMNVFLGCKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLLSVADEYDV | 242 | |
| Db | 182 | EAVDELPIINVGLFGKCVSQPEAIRQIEAGAILKIHEDWGATPMAIHCNCLNVADEMDV | 241 | |
| Qy | 243 | QVCIIHTDVNEAGYVDDTILNAMNGRAIHAYHIEGAGGGHSPDVTIMAGELNLPSTTPT | 302 | |
| Db | 242 | QVAIHSDDLNEGGFYEETVKAIAGRVIHVPHTEGAGGCHAPDVIKSVGEPNILPASTNPT | 301 | |
| Qy | 303 | IPYTTINTVAEHLDMLTCHHLDKRIREDLQPSQSRIRPGSIAAEDVLHDMGVIAMTSSDS | 362 | |
| Db | 302 | MEYTTINTVDEHLDMLVCHHLDPSPIDEDVAFAESRIIRRETTAAEDILHDMGAISVMSDS | 361 | |
| Qy | 363 | QAMGRAGEVIPTWQTADKNKCFGLPEDGKDNDFRIKRYISKYITNPALTHGVSEYI | 422 | |
| Db | 362 | QAMGRGEVVMRTWQCAHKKLQKGLSAGDTAENDNNRIKRYIAKYITNPALAHGIAHEV | 421 | |
| Qy | 423 | GSVEBGKIADLVVNNPAFGVKPKIVIKGMVVFSEMGDSNASVPTPOPVYVREMFHHG | 482 | |
| Db | 422 | GSIEKGLADIVLWDPAFGVKPKALIMKGMVAVAPMGDINAAIPTPOPVHVRPMYACIG | 481 | |
| Qy | 483 | KAKFDTSLTFVSKVAVENGVEKGLGLERQVLPVKNCRNITKKDKPFENDKTAKITVDPKTF | 542 | |
| Db | 482 | KAKYQTSMLFMSKAGIDAGVPEKGLGLSLIRGVEGCRKVTKASMIHNSYVPHIELEPQYI | 541 | |
| Qy | 543 | EYFVDEKGLCTSKPTSQVPLAQRYYTF | 568 | |


```
Db 123 ABGLIVTAGGIDTHIHFIISPOQIPAFASGITMTIMGGTGADGTNATTITPGRNWLKTM 182
QY 182 LRAAEYSNMVGLGKGNSSKQLVEQVAGAGIAGFKLHEDWGTTPSAIDHCLSVADYD 241
Db 183 LRASEYAMNLYLCKGNVSEPSLVQLEAGAGIAGFKLHEDWGSTPAIYHCLNVADKYD 242
QY 242 VQVCHTDTVNEAGYVDDTLNANMGRAIHAYHIEGAGGHSPOVITMAGELNLPSSSTP 301
Db 243 VQVAIHDTDLNAGCVEDTLQIAGRTIHTFTEGAGGHAPOVVKMSGEFNILPASTNP 302
QY 302 TIPYTIINTVAEHLDMMLTCHHLDKRIREDLQFSQSRIIRPGSIAAEDVLHDMGVIAWTS 361
Db 303 TIPFTVNTAEHMDMLVCHLHDKRIREDLQFSQSRIIRPGSIAAEDVLHDMGVIAWTS 362
QY 362 SQAMGRAGEVITRTWQTADKNKKEFGRLPEEKDNDNFRIKRYISKYTINPALTHGVSEY 421
Db 363 SQAMGRAGEVITRTWQTADKNKKEFGRLPEEKDNDNFRIKRYISKYTINPALTHGVSEY 422
QY 422 IGSVEEGKIADLVNMPAFPGVKPIVKGVMVVFSEMGDSNASVPTPOPVYVREMFHGH 481
Db 423 VGSVEVGKADLVNMPAFPGVKPIVKGVMVVFSEMGDSNASVPTPOPVYVREMFHGH 482
QY 482 GKAKEDTSITFVSKVAYENGVEKGLERQVLVKNCRNITKDFKFNDDTKAKITVDPK 541
Db 483 GKAKEDTSITFVSKVAYENGVEKGLERQVLVKNCRNITKDFKFNDDTKAKITVDPK 542
QY 542 FEVFDGKLTCTSKPTSQVPLAQRYTFF 568
Db 543 YKVKVDGKEVTSKAADKISLAQLYNLF 569
```

RESULT 2

```
QYX3W5_HELPY
ID Q7X3W5_HELPY PRELIMINARY; PRT; 569 AA.
AC Q7X3W5;
DT 01-OCT-2003 (TremBLrel. 25, Created)
DT 01-OCT-2003 (TremBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE Urease B.
GN Name=ureB;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MEL-HP27;
RA Duan G.C., Dai L.P.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY295085; AAP51176.1; -; Genomic_DNA.
DR HSSP; P14917; 1E9Y.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016151; F:nickel ion binding; IEA.
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0006807; P:nitrogen compound metabolism; IEA.
DR InterPro; IPR006680; Amidohydro.1.
DR InterPro; IPR005848; Urease_alpha.
DR InterPro; IPR008295; Urease_alphaN.
DR InterPro; IPR011612; Urease_alphaN.
DR Pfam; PF01979; Amidohydro.1; 1.
DR Pfam; PF00449; Urease_alpha; 1.
DR PIRSF; PIRSF001226; Urease_alpha; 1.
DR PRINTS; PR01752; UREASE.
DR TIGRfams; TIGR01792; urease_alpha; 1.
DR PROSITE; PS01120; UREASE_1; 1.
DR PROSITE; PS00145; UREASE_2; 1.
SQ SEQUENCE 569 AA; 61670 MW; 4C96852826C7A784 CRC64;
```

Query Match 76.7%; Score 2301; DB 2; Length 569;
Best Local Similarity 73.5%; Pred. No. 3.4e-149;
Matches 417; Conservative 75; Mismatches 75; Indels 0; Gaps 0;

QY 2 KMKKQBYVNTYGTGDKVRLGDTDLWAEVEHDYTYIGEELKFGACKTIREGMSQNSPD 61

```
Db 3 KISRKEYSVMYGPTTGDKVRLGDTDLWAEVEHDYTYIGEELKFGACKTIREGMSQNSPS 62
QY 62 ENTLDLIVNTAMIIDYTIYKADIGIKNGKIHGIGKAGNKMDDGVSPHVVGVTEALA 121
Db 63 KEELDIIITNALIVDYTIYKADIGIKNGKIHGIGKAGNKMDDGVSPHVVGVTEALA 122
QY 122 GEGMIITAGGIDSHTHFLSPQFPPTALANGVTMTMFGGGTGPDVGTNATTITPGRNWLK 181
Db 123 GEGMIITAGGIDSHTHFLSPQFPPTALANGVTMTMFGGGTGPDVGTNATTITPGRNWLK 182
QY 182 LRAAEYSNMVGLGKGNSSKQLVEQVAGAGIAGFKLHEDWGTTPSAIDHCLSVADYD 241
Db 183 LRAAEYSNMVGLGKGNSSKQLVEQVAGAGIAGFKLHEDWGTTPSAIDHCLSVADYD 242
QY 242 VQVCHTDTVNEAGYVDDTLNANMGRAIHAYHIEGAGGHSPOVITMAGELNLPSSSTP 301
Db 243 VQVAIHDTDLNAGCVEDTLQIAGRTIHTFTEGAGGHAPOVVKMSGEFNILPASTNP 302
QY 302 TIPYTIINTVAEHLDMMLTCHHLDKRIREDLQFSQSRIIRPGSIAAEDVLHDMGVIAWTS 361
Db 303 TIPFTVNTAEHMDMLVCHLHDKRIREDLQFSQSRIIRPGSIAAEDVLHDMGVIAWTS 362
QY 362 SQAMGRAGEVITRTWQTADKNKKEFGRLPEEKDNDNFRIKRYISKYTINPALTHGVSEY 421
Db 363 SQAMGRAGEVITRTWQTADKNKKEFGRLPEEKDNDNFRIKRYISKYTINPALTHGVSEY 422
QY 422 IGSVEEGKIADLVNMPAFPGVKPIVKGVMVVFSEMGDSNASVPTPOPVYVREMFHGH 481
Db 423 VGSVEVGKADLVNMPAFPGVKPIVKGVMVVFSEMGDSNASVPTPOPVYVREMFHGH 482
QY 482 GKAKEDTSITFVSKVAYENGVEKGLERQVLVKNCRNITKDFKFNDDTKAKITVDPK 541
Db 483 GKAKEDTSITFVSKVAYENGVEKGLERQVLVKNCRNITKDFKFNDDTKAKITVDPK 542
QY 542 FEVFDGKLTCTSKPTSQVPLAQRYTFF 568
Db 543 YHVFVDGKEVTSKPATKVSQAQPSIF 569
```

RESULT 3

```
URE2_HELHE
ID URE2_HELHE STANDARD; PRT; 568 AA.
AC P42823;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Urease beta subunit (EC 3.5.1.5) (Urea amidohydrolase).
GN Name=ureB;
OS Helicobacter heilmannii.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=35817;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC STRAIN=2;
RX MEDLINE=94222523; PubMed=8168924;
RA Solnick J.V., O'Rourke J., Lee A., Tompkins L.S.;
RT "Molecular analysis of urease genes from a newly identified uncultured species of Helicobacter."
RL Infect. Immun. 62:1631-1638(1994).
CC -i- CATALYTIC ACTIVITY: Urea + H(2)O = CO(2) + 2 NH(3).
CC -i- COFACTOR: Binds 2 nickel ions per subunit (By similarity).
CC -i- SUBUNIT: Consists of two subunits (alpha and beta).
CC -i- PTM: Lys-219 is carbamylated. The carbamoyl group provides the ligands for the two nickel ions (By similarity).
CC -i- SIMILARITY: Belongs to the urease family.
CC -i- CAUTION: In Helicobacter the beta subunit is what is known in other bacteria as the alpha subunit.
-----
This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its
```


use as long as its content is in no way modified and this statement is not removed.

```
CC -----
CC EMBL; L25079; AAA65723.1; -; Genomic_DNA.
CC HSSP; P14917; LE9Z.
DR InterPro; IPR006680; Amidohydro_1.
DR InterPro; IPR005848; Urease_alpha.
DR InterPro; IPR011612; Urease_alpha_N.
DR InterPro; IPR008295; Urease_alphalone.
DR Pfam; PF01979; Amidohydro_1; 1.
DR Pfam; PF00449; Urease_alpha; 1.
DR PIRSF; PIRSF001226; Urease_alpha; 1.
DR PRINTS; PR01752; UREASE.
DR TIGRFAMs; TIGR01792; urease_alpha; 1.
DR PROSITE; PS01120; UREASE_1; 1.
DR PROSITE; PS00145; UREASE_2; 1.
KW Hydrolyase; Metal-binding; Nickel.
FT ACT_SITE 321 321 Proton donor (By similarity).
FT METAL 136 136 Nickel 2 (By similarity).
FT METAL 138 138 Nickel 2 (By similarity).
FT METAL 219 219 Nickel 1 and 2 (By similarity).
FT METAL 248 248 Nickel 1 (By similarity).
FT METAL 274 274 Nickel 1 (By similarity).
FT METAL 361 361 Nickel 2 (By similarity).
FT BINDING 221 221 Substrate (By similarity).
SQ SEQUENCE 568 AA; 61870 MW; E2C036DF30377917 CRC64;

Query Match 76.6%; Score 2298.5; DB 1; Length 568;
Best Local Similarity 73.4%; Pred. No. 5e-149;
Matches 418; Conservative 74; Mismatches 74; Indels 1; Gaps 1;

QY 2 KMKKOEYVNTYGTGDKVRLGDTDLWAEVHDYTYGEEKFGAGKTIREGMQSNPSD 61
DB 3 KISRKEYVSMYGTGDKVRLGDTDLWAEVHDCTTYGEEKFGAGKTIREGMQSNPS 62

QY 62 ENTLDIVITNMIIDYTYGKADIGIKNGKIHGKAGNKMDDGVSPMVVVGTEALA 121
DB 63 SHELDIVITNALIVDYTYGKADIGIKNGKIHGKAGNKLDDGVNRLCVGPATEALA 122

QY 122 GSGMITAGDISHTHFLSPQFPPTALANGVYTMFGGGTGPVDGTTNATITPKWNLHRM 181
DB 123 AEGLIIVTAGIDITHFISFQPIPTAFASGITTMIGGTPADGTTNATITPGRWNLKEM 182

QY 182 IRAAEYSNNVGLFGKNSSSKQLVEQVAGAGIKGLHEDWGTTPSAIDHCLSVAD EYD 241
DB 183 LRASEYAMNLYGLKGNVSFPALIDQLEAGAGIKGLHEDWGSTPSAINHALNIADKYD 242

QY 242 VQVCITHTVNBAGYVDDTLNAMNGRAIHAYHIEGAGGGHSPDVTWAGELNLPSTTP 301
DB 243 VQVAIHTDNLNAGCVEDTLEAIGRTIHTFTHTGAGGGHAPDVIKWAGEFNILPASTNP 302

QY 302 TIPTYTINTVAEHLDMMLTCHLDKRIREDLQFSQRSIRPQSIAAEDVLHDMGVIAMTSSD 361
DB 303 TIPTTKNTEAEHMDMLM-CHLDKNIKEDVEFADSRIRPQTIAAEDKLHDMGIFSIITSSD 361

QY 362 SQAMGRAGEVIRPTWQTADKNKPEFKLPEDGKNDNFRIKRYISKYITNPALTHGVSEY 421
DB 362 SQAMGRAGEVIRPTWQTADKNKPEFKLPEDGKNDNFRIKRYISKYITNPALTHGISEY 421

QY 422 IGSVEGKTDADLVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPQPVYVREMGHH 481
DB 422 VGSVEGKTDADLVWNPAPFGIKPNMIIKGGFTALSQMGDANASIPTPQPVYVREMGHH 481

QY 482 GKAKFDTSITFVSKVAYENGVEKGLERQVLVPRKNCRNITKKDKFKFNDKTAITYDPKT 541
DB 482 GKAKFDTNITFVSOVAYENGKHELGLQRLVLPVKNCRNITKKDLKFNVDVTAHVENPET 541

QY 542 FEVFDVGLKLCSTSKPTSQVPLAQRYTFF 568
DB 542 YKVKVDGNEVTHAADKLSLAQLYNLF 568

RESULT 4
```

```
URE2_HELPJ STANDARD; PRT; 569 AA.
ID URE2_HELPJ
AC P69997; P14917;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Urease beta subunit (EC 3.5.1.5) (Urea amidohydrolase).
GN Name=ureB; Synonyms=hpuB; OrderedLocusNames=JHP0067;
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=99120557; PubMed=9923682; DOI=10.1038/16495;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
CC -!- CATALYTIC ACTIVITY: Urea + H(2)O = CO(2) + 2 NH(3).
CC -!- COFACTOR: Binds 2 nickel ions per subunit (By similarity).
CC -!- SUBUNIT: Heterohexamer of 3 alpha and 3 beta subunits. Four
heterohexamers assemble to form a 16 nm spherical complex (By
similarity).
CC -!- PWM: Lys-219 is carbamylated. The carbamoyl group provides the
ligands for the two nickel ions (By similarity).
CC -!- MISCELLANEOUS: Oligomerization may protect the enzyme against
denaturation in an acidic environment (By similarity).
CC -!- SIMILARITY: Belongs to the urease family.
CC -!- CAUTION: In Helicobacter the beta subunit is what is known in
other bacteria as the alpha subunit.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC -----
CC EMBL; AE001446; AAD05651.1; -; Genomic_DNA.
DR EIR; B38537; URKCBP.
DR InterPro; IPR006680; Amidohydro_1.
DR InterPro; IPR005848; Urease_alpha.
DR InterPro; IPR011612; Urease_alpha_N.
DR InterPro; IPR008295; Urease_alphalone.
DR Pfam; PF01979; Amidohydro_1; 1.
DR Pfam; PF00449; Urease_alpha; 1.
DR PIRSF; PIRSF001226; Urease_alpha; 1.
DR PRINTS; PR01752; UREASE.
DR TIGRFAMs; TIGR01792; urease_alpha; 1.
DR PROSITE; PS01120; UREASE_1; 1.
DR PROSITE; PS00145; UREASE_2; 1.
KW Complete proteome; Hydrolase; Metal-binding; Nickel.
FT ACT_SITE 322 322 Proton donor (By similarity).
FT METAL 136 136 Nickel 2 (By similarity).
FT METAL 138 138 Nickel 2 (By similarity).
FT METAL 219 219 Nickel 1 and 2 (By similarity).
FT METAL 248 248 Nickel 1 (By similarity).
FT METAL 274 274 Nickel 1 (By similarity).
FT METAL 362 362 Nickel 2 (By similarity).
FT BINDING 221 221 Substrate (By similarity).
SQ SEQUENCE 569 AA; 61684 MW; 4C8A6BC6C8295584 CRC64;

Query Match 76.6%; Score 2297; DB 1; Length 569;
Best Local Similarity 73.4%; Pred. No. 6.3e-149;
Matches 416; Conservative 76; Mismatches 75; Indels 0; Gaps 0;

QY 2 KMKKOEYVNTYGTGDKVRLGDTDLWAEVHDYTYGEEKFGAGKTIREGMQSNPSD 61
DB 3 KISRKEYVSMYGTGDKVRLGDTDLWAEVHDYTYGEEKFGAGKTIREGMQSNPS 62
```



```
Db 362 QAMGRVGEVITRTWQTADKNKKEFGRLKEEKGDNDNFRIKRYLSKYTINPAIAHGISEYV 421
QY 423 GSVEEKGADLVVNPAFFGVKPKIVIKGGMVVFSEMGDSNASVPTPOPYYREMFHGHG 482
Db 422 GSVEGKADLVVNSPAFFGVKPNMIIKGFIALSQMGDANASIPTPOPYYREMFHGHG 481
QY 483 KAKFDTSIITFVSKVAYENGVEKELGLERQVLPVKNCRNITKDKFNDKTAKITVDPKTF 542
Db 482 KAKYDANITFVSOAAYDKGKIEELGLERQVLPVKNCRNITKDKMFNDTTAHIEVNPETY 541
QY 543 EYFVVGKCTSKPTSQVPLAQRVTF 568
Db 542 HVFVVGKVTSPATKVSQAQFSIF 567

RESULT 8
Q9AFBI_HELPY
ID Q9AFBI_HELPY PRELIMINARY; PRT; 569 AA.
AC Q9AFBI;
DT 01-JUN-2001 (T-EMBLrel. 17, Created)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Urease B.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Gerold A., Simpson K., Chang Y.-F.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF352376; AAK32714.1; -; Genomic_DNA.
DR HSSP; P14917; IE9Y.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016151; F:nickel ion binding; IEA.
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0006807; P:nitrogen compound metabolism; IEA.
DR InterPro; IPR006680; Amidohydro_1.
DR InterPro; IPR005848; Urease alpha.
DR InterPro; IPR008295; Urease alphas.
DR InterPro; IPR011612; Urease alpha_N.
DR Pfam; PF00449; Urease_alpha; 1.
DR PIRSF; PIRSF001226; Urease_alpha; 1.
DR PRINTS; PR01752; UREASE.
DR TIGRFAMs; TIGR01792; urease_alph; 1.
DR PROSITE; PS01120; UREASE_1; 1.
DR PROSITE; PS00145; UREASE_2; 1.
SQ SEQUENCE 569 AA; 61713 MW; C243ABC38C25D187 CRC64;

Query Match 76.4%; Score 2291; DB 2; Length 569;
Best Local Similarity 73.2%; Pred. No. 1.6e-148;
Matches 415; Conservative 76; Mismatches 0; Indels 0; Gaps 0;

QY 2 KMKKQEVNTYGTGDKVRLGDTDLWAEVEHDYTYGEEELKFGAGKTIREGMQSNPSD 61
Db 3 KISRKEYVMYGTGDKVRLGDTDLWAEVEHDYTYGEEELKFGAGKTIREGMQSNPS 62
QY 62 ENTLDLIVTNAMITDYGKADIGIKNGKIHGKAGNKMDDGVSPHVMVGVGTEALA 121
Db 63 KEELDLIITNALIVDYGKADIGIKNGKIHGKAGNKMDDGVSNVNSVGPATEALA 122

QY 122 GEGMIITAGGIDSHFLSPQPPPTALANGVTMTFGGGTGPVDTGNATTTTPGKNLHRM 181
Db 123 GEGLIIVTAGGIDTHIFISPPQIPTAFASGVTMTWGGTGPADGTNATTTTPGRNLKM 182
QY 182 LRAEYYSNNVGLFGKNSSSKQVVEAGIAGFKLHEDWGTTPSAIDHCLSVADEYD 241
Db 183 LRAAEYYSNNLSPLAGNASNDASLADQIEAGAIGFKIHEDWGTTPSAINHDLVDADKYD 242
QY 242 VQVCIHDTVNEAGYVDDTLNANNGRAIHAYHIEGAGGHSVDVITWAGELNITLPSSTP 301
Db 243 VQVAIHDTDLNEAGCVEDTMAAAGRMTWHTFTTEGAGGHPADIIKVAGEHNILPASTNP 302
QY 302 TIPTYTINTVAEHLDMLTCHHLDRKIREDLQFSQSRIRPGSIAAEDVLHDMGVIAWTSSD 361
Db 303 TIPTVTATEAHEMDMLMVCHLDKSIKEDVQFADSRIRPQTIAAEDTLHDMGIFSISSD 362
QY 362 SQAMGRAGEVITRTWQTADKNKKEFGKLPEDGKUNDNFRIKRYLSKYTINPAIAHGISEY 421
Db 363 SQAMGRVGEVITRTWQTADKNKKEFGRLKEEKGDNDNFRIKRYLSKYTINPAIAHGISEY 422
QY 422 IGSVEEKGADLVVNPAFFGVKPKIVIKGGMVVFSEMGDSNASVPTPOPYYREMFHGHG 481
Db 423 VGSVEGKADLVVNSPAFFGVKPNMIIKGFIALSQMGDANASIPTPOPYYREMFHGHG 482
QY 482 GKAKFDTSIITFVSKVAYENGVEKELGLERQVLPVKNCRNITKDKFNDKTAKITVDPKTF 541
Db 483 GKAKYDANITFVSOAAYDKGKIEELGLERQVLPVKNCRNITKDKMFNDTTAHIEVNPETY 542
QY 542 FEVFDGKCTSKPTSQVPLAQRVTF 568
Db 543 YHVFVVGKVTSPATKVSQAQFSIF 569

RESULT 9
Q84F75_HELPY
ID Q84F75_HELPY PRELIMINARY; PRT; 569 AA.
AC Q84F75;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Urease beta (EC 3.5.1.5).
GN Name=ureB;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Karimi Arzenani M., Mohammadi M.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY227442; AAO34403.1; -; Genomic_DNA.
DR HSSP; P14917; IE9Y.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016151; F:nickel ion binding; IEA.
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0006807; P:nitrogen compound metabolism; IEA.
DR InterPro; IPR005848; Urease alpha.
DR InterPro; IPR008295; Urease alphas.
DR InterPro; IPR011612; Urease_alpha_N.
DR Pfam; PF01979; Amidohydro_1; 1.
DR Pfam; PF00449; Urease_alpha; 1.
DR PIRSF; PIRSF001226; Urease_alpha; 1.
DR PRINTS; PR01752; UREASE.
DR TIGRFAMs; TIGR01792; urease_alph; 1.
DR PROSITE; PS01120; UREASE_1; 1.
DR PROSITE; PS00145; UREASE_2; 1.
KW Hydrolase.
SQ SEQUENCE 569 AA; 61697 MW; D35279ED45EF712F CRC64;

Query Match 76.3%; Score 2288; DB 2; Length 569;
Best Local Similarity 73.2%; Pred. No. 2.6e-148;
Matches 415; Conservative 75; Mismatches 77; Indels 0; Gaps 0;

QY 2 KMKKQEVNTYGTGDKVRLGDTDLWAEVEHDYTYGEEELKFGAGKTIREGMQSNPSD 61
Db 3 KISRKEYVMYGTGDKVRLGDTDLWAEVEHDYTYGEEELKFGAGKTIREGMQSNPS 62
QY 62 ENTLDLIVTNAMITDYGKADIGIKNGKIHGKAGNKMDDGVSPHVMVGVGTEALA 121
Db 63 KEELDLIITNALIVDYGKADIGIKNGKIHGKAGNKMDDGVSNVNSVGPATEALA 122
```

QY 122 GEGMIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGGTGPVDTGNATTITPGKWNLRHM 181
Db 123 GEGLVITAGGIDTHIHFIHQIIPAFASGVTTMIGGGTGPADGTTNATTITPGRRLKWM 182
QY 182 LRAAEYSNNVGLGKGNSSKKQLVEQVEAGAIGFKLHEDGTTTPSAIDHCLSVADYD 241
Db 183 LRAAEYSNNLGLAKGNASNDASLADQIEAGAIGFKIHEDGTTTPSAINHALDVADKYD 242
QY 242 VQVCIHDTVNEAGYVDDTLNMGNGRAIHAYHIEGAGGGHSPDVITMAGELNLPSTTP 301
Db 243 VQVAIHDTLNEAGCVKDTMAALAGRTMTHTFTEGAGGGHAPDIIKVAGEHNILPASTNP 302
QY 302 TTPYTIINTVAEHLDMMLTCHHLDKRIREDLQFSQSRIRPGSIAAEVLDHDMGVIAMTSSD 361
Db 303 TTPFTVNPBAEHWMLMVCHLDKSIKEDVQFADSRIRPQTIAAEDTLHDMGIFSITSSD 362
QY 362 SQAMGRAGEVITRTWTQADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPAITHGVSEY 421
Db 363 SQAMGRVGEVITRTWTQADKNKKEFGRLKEEGKNDNFRIKRYLSKYITINPAIAHIGSEY 422
QY 422 IGSVEGKIADLVVWNPAPFVGKPKIVIKGGMVVFSEMGDSNASVPTPOPVYVREMFGHH 481
Db 423 VGSVEGVKADLVVWNPAPFVGKPKMIIKGGFIALSQMGDANASITPTPOPVYVREMFAHH 482
QY 482 GKAKEDTSTTFVSKVAYENGVEKGLERQVLVVKNCRNITKDFKFNDKTAKITVDPKT 541
Db 483 GKAKYDANITFVSOAAYDKGIKEELGLERQVLVVKNCRNITKDMQFNDTTAHIEVNPET 542
QY 542 FEVFDGKLTCKTSKPSQVLAQRYTF 568
Db 543 YRVFVDGKEVTSKPANKVSLAQLSIF 569
RESULT 10
Q8RNU6_HELPY PRELIMINARY; PRT; 559 AA.
ID Q8RNU6_HELPY
AC Q8RNU6_HELPY
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Urease B subunit.
GN Name=ureB;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CH-CTX1;
RX MEDLINE=20123532; PubMed=10660136;
RA Opazo P., Muller I., Rolian A., Valenzuela P., Yudelevich A.,
Garcia-de la Guardia R., Urra S., Venegas A.;
RT "Serological response to Helicobacter pylori recombinant antigens in
Chilean infected patients with duodenal ulcer, non-ulcer dyspepsia and
gastric cancer.";
RL APMIS 107:1069-1078 (1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CH-CTX1;
RX MEDLINE=22121498; PubMed=12125208;
RA Muller I., Medina-Selby A., Palacios J.L., Martinez P., Opazo P.,
Bruce E., Mancilla M., Valenzuela P., Yudelevich A., Venegas A.;
RT "Cloning and comparison of ten gene sequences of a Chilean H. pylori
strain with other H. pylori strains revealed higher variability for
VacA and CagA virulence factors.";
RL Biol. Res. 35:67-84 (2002).
DR EMBL: AF479026; AAL86896.1; -; Genomic_DNA.
DR HSSP: P14917; 1E9Y.
DR SMR; Q8RNU6; 1-553.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016151; F:nickel ion binding; IEA.
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0006807; P:nitrogen compound metabolism; IEA.

DR InterPro; IPR006680; Amidohydro_1.
DR InterPro; IPR005848; Urease_alpha.
DR InterPro; IPR008295; Urease_alphalone.
DR InterPro; IPR011612; Urease_alpha_N.
DR Pfam; PF01979; Amidohydro_1_1.
DR Pfam; PF00449; Urease_alpha; 1.
DR PRF01; PRF001226; Urease_alpha; 1.
DR PRF01; PRF01752; UREASE.
DR TIGRFAMs; TIGR01792; urease_alph; 1.
DR PROSITE; PS01120; UREASE_1; 1.
DR PROSITE; PS00145; UREASE_2; 1.
SQ SEQUENCE 559 AA; 60607 MW; 049CA0F87E2CABID CRC64;
Query Match 75.1%; Score 2251; DB 2; Length 559;
Best Local Similarity 73.9%; Pred. No. 8.8e-146;
Matches 407; Conservative 72; Mismatches 72; Indels 0; Gaps 0;
QY 2 KMKKQEVNTYGTGDKVRLGDTDLWAEVHDYTTTGEBELKFGAGKTIRGCMQSNPSPD 61
Db 3 KISRKEYASMYGTTGDKVRLGDTDLJAEVHDYTYGEBELKFGGKTLRGMSQSNPS 62
QY 62 ENTLDLVIITNAMIIDYTYGIKADIGIKNGKHIGIKAGNKDMQDGVSPHVVGVCTEALA 121
Db 63 KEELDUITINALLVDYTYGIKADIGIKNGKHIGIKAGNKDMQDGVKNLSVGPATEALA 122
QY 122 GEGMIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGGTGPVDTGNATTITPGKWNLRHM 181
Db 123 GEGLVITAGGIDTHIHFIHQIIPAFASGVTTMIGGGTGPADGTTNATTITPGRRLKWM 182
QY 182 LRAAEYSNNVGLGKGNSSKKQLVEQVEAGAIGFKLHEDGTTTPSAIDHCLSVADYD 241
Db 183 LRAAEYSNNLGLAKGNASNDASLADQIEAGAIGFKIHEDGTTTPSAINHALDVADKYD 242
QY 242 VQVCIHDTVNEAGYVDDTLNMGNGRAIHAYHIEGAGGGHSPDVITMAGELNLPSTTP 301
Db 243 VQVAIHDTLNEAGCVKDTMAALAGRTMTHTFTEGAGGGHAPDIIKVAGEHNILPASTNP 302
QY 302 TTPYTIINTVAEHLDMMLTCHHLDKRIREDLQFSQSRIRPGSIAAEVLDHDMGVIAMTSSD 361
Db 303 TTPFTVNPBAEHWMLMVCHLDKSIKEDVQFADSRIRPQTIAAEDTLHDMGIFSITSSD 362
QY 362 SQAMGRAGEVITRTWTQADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPAITHGVSEY 421
Db 363 SQAMGRVGEVITRTWTQADKNKKEFGRLKEEGKNDNFRIKRYLSKYITINPAIAHIGSEY 422
QY 422 IGSVEGKIADLVVWNPAPFVGKPKIVIKGGMVVFSEMGDSNASVPTPOPVYVREMFGHH 481
Db 423 VGSVEGVKADLVVWNPAPFVGKPKMIIKGGFIALSQMGDANASITPTPOPVYVREMFAHH 482
QY 482 GKAKEDTSTTFVSKVAYENGVEKGLERQVLVVKNCRNITKDFKFNDKTAKITVDPKT 541
Db 483 GKAKYDANITFVSOAAYDKGIKEELGLERQVLVVKNCRNITKDMQFNDTTAHIEVNPET 542
QY 542 FEVFDGKLTCKTSKPSQVLAQRYTF 552
Db 543 YRVFVDGKEVT 553
RESULT 11
URE2_HELPY
ID URE2_HELPY STANDARD; PRT; 569 AA.
AC Q08716; Q08716;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Urease beta subunit (EC 3.5.1.5) (Urea amidohydrolase).
GN Name=ureB;
OS Helicobacter felis.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=214;
RN [1]
RP NUCLEOTIDE SEQUENCE.

```

RX MEDLINE=94018627; PubMed=8412683;
RA Ferrero R.L., Labigne A.;
RT "Cloning, expression and sequencing of Helicobacter felis urease
RT genes.";
RL Mol. Microbiol. 9:323-333(1993).
RN [2]
RP PROTEIN SEQUENCE OF 1-12.
RC STRAIN=ATCC 49179;
RX PubMed=1452359;
RA Turbett G.R., Hoef P.B., Horne R., Mee B.J.;
RT "Purification and characterization of the urease enzymes of
RT Helicobacter species from humans and animals.";
RL Infect. Immun. 60:5259-5266(1992).
CC -!- CATALYTIC ACTIVITY: Urea + H(2)O = CO(2) + 2 NH(3).
CC -!- COFACTOR: Binds 2 nickel ions per subunit (By similarity).
CC -!- SUBUNIT: Consists of two subunits (alpha and beta).
CC -!- PTM: Lys-219 is carbamylated. The carbamoyl group provides the
CC ligands for the two nickel ions (By similarity).
CC -!- SIMILARITY: Belongs to the urease family.
CC -!- CAUTION: In Helicobacter the beta subunit is what is known in
CC other bacteria as the alpha subunit.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; X69080; CA448826.1; -; Genomic_DNA.
DR PIR; C49215; C49215.
DR PIR; S35291; S35291.
DR HSSP; P14917; 1E9Z.
DR InterPro; IPRO06680; Amidohydro 1.
DR InterPro; IPRO05848; Urease_alpha.
DR InterPro; IPRO11612; Urease_alpha_N.
DR InterPro; IPRO08295; Urease_alphalone.
DR Pfam; PF01979; Amidohydro 1; 1.
DR Pfam; PF00449; Urease_alpha; 1.
DR PIRSF; PIRSF001226; Urease_alpha; 1.
DR PRINTS; PR01752; UREASE.
DR TIGRFAMs; TIGR01792; urease_alph; 1.
DR PROSITE; PS01120; UREASE 1; 1.
DR PROSITE; PS00145; UREASE 2; 1.
KW Direct protein sequencing; Hydrolase; Metal-binding; Nickel.
FT ACT_SITE 322 322 Proton donor (By similarity).
FT METAL 136 136 Nickel 2 (By similarity).
FT METAL 138 138 Nickel 2 (By similarity).
FT METAL 219 219 Nickel 1 and 2 (By similarity).
FT METAL 248 248 Nickel 1 (By similarity).
FT METAL 274 274 Nickel 1 (By similarity).
FT METAL 362 362 Nickel 2 (By similarity).
FT BINDING 221 221 Substrate (By similarity).
SQ SEQUENCE 569 AA; 61703 MW; 66DBB9952CC7CEA3 CRC64;

Query Match 74.8%; Score 2244; DB 1; Length 569;
Best Local Similarity 72.5%; Pred. No. 2.7e-145;
Matches 411; Conservative 71; Mismatches 85; Indels 0; Gaps 0;

QY 2 KMKKQYVNTYGTGDKVRLGDTDLWAEVHDYTTYGEELKFGAGKTIREGNGQNSPD 61
DB 3 KISRKEYSVSYGPTTGDVRLGDTDLLEVEHDTTYGEEIKFGGKTIIRDGMSQNSPS 62
QY 62 ENTLDLVITNAMITDVTGYIKADIGIKNGKIHGIGKAGNKDMQGVSPHVVGVGTEALA 121
DB 63 SYELDLVLNALIVDVTGYIKADIGIKDGKGIAGIKAGNKDMQGVNDNLVCGVPATEALA 122
QY 122 GEGMIITAGGIDSHTHFLSPQPTALANGVTTFMFGGTGTPVGTNATITPCKNLHRM 181
DB 123 AEGLIIVTAGGIDTHIHFIISQPIPTAFASGVTTFMIGGTGPADGTNATITPGRANLKSM 182
QY 182 LRAAEYSNMVGLFGKNSSSKKQLEQVEQVAGAIQFKLHEDWGTTPSAIDHCLSVADEYD 241
DB 183 LRAAEYAMNLGLFAGKNSVYFSLRDQIEAGAIQFKIHEDWGSTPAIIHCLINVADEYD 242

```

```

QY 242 VQVCIHDTTNEAGYVDVDTLNMNGRAIHAYHIEGAGGSHSPDVTMAGELNILPSSTTP 301
DB 243 VQVAIHTDTLNEAGCVEDTLEAIGARTIHTFHTGAGGSHAPDVVKMAGEFNILPASTNP 302
QY 302 TIPYTINTVAEHLDMLTCHHLDKRIEDLOFSQSRIIRPGSIAAEDVLHDMGVIAWTSSD 361
DB 303 TIPFTKTEAEHMDMLMVCCHLDKSIKEDVOFADSRIRPQTIAAEDQLHDMGIFSTSSD 362
QY 362 SQAMGRAGEVTPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTHTGVSEY 421
DB 363 SQAMGRGEVITRTWQTADKNKKEFGRLKEEGKNDNFRIKRYISKYITINPGIAHGISDY 422
QY 422 IGSVEGKIADLVVYVNPAPFGVKPKIVIKGMMVFESEMGDSNASVPTQPVTYIREMFGHH 481
DB 423 VGSVEVGKYADLVLSWPAFFGIKPMNIKGGFIALSQMGDANASITPTQPVTYIREMFGHH 482
QY 482 GKAKFDTSITFVSKVAYENGVEKELGLERQVLVKNCCNITKDKPKENDTKAKITVDPKT 541
DB 483 GKKNFDTNITFVSOAYKAGIKBEILGLDRAAPPVKNCRNITKKOLKFNVDVTAHIDVNPET 542
QY 542 FEVFDGKLTCKTSPSOVPLAQRVTFP 568
DB 543 YKVKVDGKEVTSKAADELSLAQLNLF 569

RESULT 12
Q93PJ4_HELHP PRELIMINARY; PRT; 569 AA.
AC Q93PJ4; Q7BYA2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Urease beta subunit UreB (EC 3.5.1.5).
GN Name=ureb; OrderedLocNames=RH0408;
OS Helicobacter hepaticus.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=32025;
RN [1]
RP NUCLEOTIDE SEQUENCE..
RX MEDLINE=21391860; PubMed=11500473;
RX DOI=10.1128/JAI.69.9.5914-5920.2001;
RA Beckwith C.S., McGee D.J., Mobley H.L.T., Riley L.K.;
RT "Cloning, expression, and catalytic activity of Helicobacter hepaticus
RT urease.";
RL Infect. Immun. 69:5914-5920(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 51449 / 3B1;
RX MEDLINE=22709201; PubMed=12810954; DOI=10.1073/pnas.1332093100;
RA Suerbaum S., Josenhans C., Sterzenbach T., Drescher B., Brandt P.,
RA Bell M., Droege M., Fartmann B., Fischer H.-P., Ge Z., Hoerster A.,
RA Holland R., Klein K., Koenig J., Macko L., Mendz G.L., Nyakatura G.,
RA Schauer D.B., Shen Z., Weber J., Frosch M., Fox J.G.;
RT "The complete genome sequence of the carcinogenic bacterium
RT Helicobacter hepaticus.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7901-7906(2003).
DR EMBL; AF332656; AAK69199.1; -; Genomic DNA.
DR EMBL; AE017145; AAP7005.1; -; Genomic DNA.
DR HSSP; P14917; 1E9Y.
DR HSSP; P18314; 1A5N.
DR SMR; Q93PJ4; 1-569.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016151; F:nickel ion binding; IEA.
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0006807; P:nitrogen compound metabolism; IEA.
DR InterPro; IPRO06580; Amidohydro 1.
DR InterPro; IPRO05848; Urease_alpha.
DR Pfam; PF01979; Amidohydro 1; 1.
DR Pfam; PF00449; Urease_alpha; 1.
DR PRINTS; PR01752; UREASE.

```



```
QY 482 GKAKEDTSITFVSKVAYENGVEKLGLEROVLVVKNCNRTTKKDFKFNKTKAKITVDPKPT 541
Db 483 GKAKYQTSITFVSKVAYENGVEKLGLEROVLVVKNCNRTTKKDFKFNKTKAKITVDPKPT 542
QY 542 FEVFDGKLCSTKPTSQVPLAQRYYTF 568
Db 543 YEVKVDGQLVTCPEAIVPMAQRYYFLF 569

RESULT 14
Q9KG59 BACHD
ID Q9KG59_BACHD PRELIMINARY; PRT; 571 AA.
AC Q9KG59;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Urease alpha subunit (EC 3.5.1.5).
GN Name=ureC; OrderedLocusNames=BH0254;
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132; DOI=10.1093/nar/28.21.4317;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331 (2000).
DR EMBL; BA000004; BAB03973.1; -; Genomic_DNA.
DR PIR; F83681; F83681.
DR HSSP; P41020; 1UBP.
DR SMR; Q9KG59; 1-571.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016151; F:nickel ion binding; IEA.
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0019627; P:urea metabolism; IEA.
DR InterPro; IPR006680; Amidohydro_1.
DR InterPro; IPR005848; Urease_alpha.
DR InterPro; IPR008295; Urease_alphalone.
DR InterPro; IPR011612; Urease_alpha_N.
DR Pfam; PF01979; Amidohydro_1; 1.
DR Pfam; PF00449; Urease_alpha; 1.
DR PIRSF; PIRSF001226; Urease_alpha; 1.
DR PRINTS; PR01752; UREASE.
DR TIGRFAMs; TIGR01792; urease_alpha; 1.
DR PROSITE; PS01120; UREASE_1; 1.
DR PROSITE; PS00145; UREASE_2; 1.
KW Complete proteome; Hydrolase.
SQ Complete 571 AA; 61291 MW; 397398DECEA0A80F CRC64;

Query Match 67.2%; Score 2016.5; DB 2; Length 571;
Best Local Similarity 64.4%; Pred. No. 1.1e-129;
Matches 368; Conservative 88; Mismatches 112; Indels 3; Gaps 2;

QY 1 MKNKKQEVNTYPTGDKVRLGDTDLMAVEHDYTYGELKFGAGKTIREGMGQS--N 58
Db 1 MKLTRAQHASLYPTGDKVRLADTOLLLEIEKDYTVYGDCKFGGKVLRDGMGQSAVY 60
QY 59 SPDENTLDLVTNMIIDYTVGIVKADIGIKNGKIHGIGKAGNKMDGVSPhMVVGVTGTE 118
Db 61 TRDEGVLDLIITNATIIDYTVGIVKADIGIKDGHVIGIKGNGNPDIIMDGVESHMIVGASTE 120
QY 119 ALAGEGMIITAGGIDSHFTLSPQFPPTALANGVTWTFGGGTGVPDGTNATITTPGKNWL 178
Db 121 ATAGEGLIVTAGGIDAHIFISFQQIDVAIASITMLGGTGTPATGTTCTPGKNNI 180
QY 179 HRMLRAAEYSVMNVLFGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVAD 238
Db 181 ERMLEAADAPPVNLGLFGKGNASTPAPRQIEAGAILGLKHLHEDWGTTPAIRTCLSVAD 240
```

```
QY 239 EYDVQVCITHTDVNEAGVYDDTLNAMNGRAIHAYHIEGAGGGHSPDVTIMAGELNLPSS 298
Db 241 RMDVQVALHTDLTNEAGFVEDTIIKAIGDRVHTYHTEGAGGGAHPDMKVAGLPNVLPS 300
QY 299 TTPTIPYTIINTVAEHLDMATCHHLDKRIRDLQFSOSRIRPGSIAAEDVLHDMGVAMT 358
Db 301 TNPTPTVTNTIDHLDMLAVCHLDPNVPEDVAFADSKIRPETIAAEDILQDLGVISMI 360
QY 359 SSDSQAMGRAGEVTPRTWTQADKNKKGKLPED-GKNDNPNRIKRYISKYINPALTHG 417
Db 361 SSDSQAMGRVGEVITRTWTQADKNKKGKALAEQDQKGNNDNRIKRYISKYINPAITHG 420
QY 418 VSEYIGSVVEGKIADLVVNPAFFGKPKIVIKGGMVVFSEMGSDSNASVPTPPVYREM 477
Db 421 IDDVYGVSEVGEKADLVLDLWDRFFGKPELILKGLIAYSQMGDPNASIPTQPVSFRPM 480
QY 478 FGHGKAKFDTSIITFVSKVAYENGVEKLGLEROVLVVKNCNRTTKKDFKFNKTKAKITV 537
Db 481 FGSFGRAGTTSITFSLKAAMDVLGVHEALGLQKIAHVKNCSISKRSKMYNDATPNIEI 540
QY 538 DPKTFFVFDGKLCSTKPTSQVPLAQRYYTF 568
Db 541 DPETVEVKVDGEMITCKPFEVALAQRYYFLF 571

RESULT 15
URE1 BACSU
ID_URE1 BACSU STANDARD; PRT; 569 AA.
AC P77837;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Urease alpha subunit (EC 3.5.1.5) (Urea amidohydrolase).
GN Name=ureC; OrderedLocusNames=BSU36640;
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=168;
RA Glaser P., Cruz H., Danchin A.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.-K., Codani J.-J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Duesterhoeft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.-D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.-Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holtsappel S., Hosono S., Hullo M.-F., Itaya M.,
RA Jones L.-M., Joris B., Karamata D., Kasahara Y., Klaer-Blanchard M.,
RA Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krogh S.,
RA Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J.,
RA Lazarevic V., Lee S.-M., Levine A., Liu H., Masuda S., Maue C.,
RA Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S.,
RA Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,
RA Park S.-H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,
RA Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G.,
RA Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B.,
RA Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S., Schroeter R.,
RA Scoffone P., Seguchi J., Sekowska A., Seror S.J., Serror P.,
RA Shin B.-S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,
RA Tognoni K., Tosechi M., Tamakoshi A., Tanaka T., Terpstra P.,
RA Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier P.,
RA Vaasarotti A., Viari A., Wambutt R., Wedler E., Wedler H.,
RA Weitzenegger T., Winters P., Wipat A., Yamamoto H., Yamane K.,
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 28, 2005, 07:55:02 ; Search time 185.28 Seconds
(without alignments)
1346.976 Million cell updates/sec

Title: US-09-904-994B-3
Perfect score: 2999
Sequence: 1 MRMKQYVNTYPTKGDV.....KLCTSKPTQVPLAQRYYTF 568

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : A_Geneseq_21.*
- 1: Geneseq1980s.*
 - 2: Geneseq1990s.*
 - 3: Geneseq2000s.*
 - 4: Geneseq2001s.*
 - 5: Geneseq2002s.*
 - 6: Geneseq2003as.*
 - 7: Geneseq2003bs.*
 - 8: Geneseq2004s.*
 - 9: Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|--------------------|
| 1 | 2999 | 100.0 | 568 | 5 | ADJ58239 Urease su |
| 2 | 2973 | 99.1 | 568 | 5 | Adj58251 UreaseXY |
| 3 | 2960 | 98.7 | 568 | 5 | Adj58242 UreaseXY |
| 4 | 2953 | 98.5 | 568 | 5 | Adj58248 UreaseXY |
| 5 | 2598 | 86.3 | 496 | 5 | ADJ58245 UreaseXY |
| 6 | 2326 | 77.6 | 569 | 8 | Adj37851 H. bizoz |
| 7 | 2297 | 76.6 | 569 | 2 | AAR04580 Part of p |
| 8 | 2297 | 76.6 | 569 | 2 | Aaw07194 H. pylori |
| 9 | 2297 | 76.6 | 569 | 2 | AAW98511 H. pylori |
| 10 | 2297 | 76.6 | 569 | 6 | ABU30687 Protein e |
| 11 | 2297 | 76.6 | 569 | 8 | ADU05536 H. pylori |
| 12 | 2282 | 76.1 | 568 | 8 | ADS09177 H. pylori |
| 13 | 2281 | 76.1 | 569 | 2 | AAR67378 H. pylori |
| 14 | 2281 | 76.1 | 570 | 4 | AEBS5116 Hylicobac |
| 15 | 2244 | 74.8 | 569 | 2 | AAR74337 Helicobac |
| 16 | 2244 | 74.8 | 569 | 2 | Aaw06730 H. felis |
| 17 | 2238 | 74.6 | 749 | 2 | AAW80599 Helicobac |
| 18 | 2237 | 74.6 | 806 | 2 | AAR67371 Urease A |
| 19 | 2233 | 74.5 | 559 | 2 | AAR13550 B subunit |
| 20 | 2225 | 74.2 | 559 | 8 | ADM28644 Helicobac |
| 21 | 2013.5 | 67.1 | 568 | 2 | AAR36387 Urease ga |
| 22 | 1922 | 64.1 | 568 | 6 | ABU19714 Protein e |
| 23 | 1914 | 63.8 | 566 | 6 | ABU41362 Protein e |
| 24 | 1907 | 63.6 | 568 | 6 | ABU21842 Protein e |

| | | | | | | |
|----|--------|------|-----|---|----------|--------------------|
| 25 | 1897 | 63.3 | 568 | 6 | ABU22314 | Abu22314 Protein e |
| 26 | 1832 | 61.1 | 566 | 6 | ABU38819 | Abu38819 Protein e |
| 27 | 1828 | 61.0 | 566 | 7 | ABO78141 | ABO78141 Pseudomon |
| 28 | 1826 | 60.9 | 572 | 6 | ABU30286 | Abu30286 Protein e |
| 29 | 1825 | 60.9 | 567 | 6 | ABU32083 | Abu32083 Protein e |
| 30 | 1821 | 60.7 | 625 | 7 | ABO62699 | ABO62699 Klebsiell |
| 31 | 1818.5 | 60.6 | 569 | 7 | ADF05744 | Adf05744 Bacterial |
| 32 | 1818.5 | 60.6 | 779 | 6 | ABU41210 | Abu41210 Protein e |
| 33 | 1818 | 60.6 | 567 | 2 | ADA34625 | Ada34625 Acinetoba |
| 34 | 1817 | 60.6 | 567 | 2 | AAW37774 | Aaw37774 Klebsiell |
| 35 | 1817 | 60.6 | 567 | 3 | AAV81825 | Aav81825 Klebsiell |
| 36 | 1815 | 60.5 | 566 | 6 | ABU17040 | Abu17040 Protein e |
| 37 | 1811.5 | 60.4 | 568 | 2 | AAK67379 | Aar67379 P. mirabi |
| 38 | 1809 | 60.3 | 568 | 6 | ABU28899 | Abu28899 Protein e |
| 39 | 1809 | 60.3 | 568 | 6 | ABU28920 | Abu28920 Protein e |
| 40 | 1809 | 60.3 | 568 | 7 | ADC00113 | Adc00113 Enterohae |
| 41 | 1797 | 59.9 | 567 | 6 | ABU40328 | Abu40328 Protein e |
| 42 | 1789.5 | 59.7 | 571 | 4 | AAK81512 | Aag81512 S. epider |
| 43 | 1789.5 | 59.7 | 811 | 6 | ABU42650 | Abu42650 Protein e |
| 44 | 1788.5 | 59.6 | 573 | 5 | ABP40181 | Abp40181 Staphyloc |
| 45 | 1788.5 | 59.6 | 573 | 8 | ADS07231 | Ads07231 Staphyloc |

ALIGNMENTS

RESULT 1
ADJ58239
ID ADJ58239 standard; protein; 568 AA.
XX
AC ADJ58239;
XX
XX
DT 06-MAY-2004 (first entry)
XX
DE Urease subunit polypeptide Y.
XX
KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.
XX
OS Helicobacter felis.
XX
PN EPI176192-A2.
XX
PD 30-JAN-2002.
XX
PF 11-JUL-2001; 2001EP-00202666.
XX
PR 17-JUL-2000; 2000EP-00202565.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Kusters JG, Cattoli G;
XX
DR WPI; 2002-124384/17.
XX
DR N-PSDB; ADJ58237.
XX
PT Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
the diagnosis of Helicobacter felis infections and in the preparation of
vaccines.
XX
PS Claim 10; SEQ ID NO 3; 76pp; English.
XX
CC The present invention relates to a novel Helicobacter felis urease X and
Y subunit polypeptides and immunogenic fragments. The polypeptides are
used in the manufacture of vaccines against Helicobacter felis infections
and in diagnostic tests to detect antibodies against Helicobacter felis.
CC
CC Helicobacter felis is difficult to grow so it is more convenient to use
the expression products of the genes encoding the urease X and Y subunits
in the manufacture of vaccines. The present sequence represents an urease
X subunit polypeptide of the invention.
XX
SQ Sequence 568 AA;

Query Match 100.0%; Score 2999; DB 5; Length 568;


```
QY 241 DVQVCIHDTVNEAGVDDTLNAMNGRAIHAYHIEGAGGHSPPDVITMAGELNILPSSTT 300
Db 241 DVQVCIHDTVNEAGVDDTLNAMNGRAIHAYHIEGAGGHSPPDVITMAGELNILPSSTT 300
QY 301 PTPIPTINTVAEHLDMLTCHLDKRIREDLOFSQSRIRPGSIAAEDVLHDMGVIAWTSS 360
Db 301 PTPIPTINTVAEHLDMLTCHLDKRIREDLOFSQSRIRPGSIAAEDVLHDMGVIAWTSS 360
QY 361 DSQAMGRAGEVIPRTWQTADKNKCFGLPEQKNDNFRIKRYISKYTTINPALTHGVSE 420
Db 361 DSQAMGRAGEVIPRTWQTADKNKCFGLPEQKNDNFRIKRYISKYTTINPALTHGVSE 420
QY 421 YIGSVEEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPQPVYIREMFGH 480
Db 421 YIGSVEEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPQPVYIREMFGH 480
QY 481 HGKAKFDTSTFVSKVAYENGVEKLGLEROVLVKNCRNITKDPKFNKTAKITVDPK 540
Db 481 HGKAKFDTSTFVSKVAYENGVEKLGLEROVLVKNCRNITKDPKFNKTAKITVDPK 540
QY 541 TFEVFDGKLCSTKPSQVPLAQRYTF 568
Db 541 TFEVFDGKLCSTKPSQVPLAQRYTF 568

RESULT 5
ADJ58245
ID ADJ58245 standard; protein; 496 AA.
XX AC ADJ58245;
XX DT 06-MAY-2004 (first entry)
XX DE UreaseXY subunit #4.
XX KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.
XX OS Helicobacter felis.
XX PN EP1176192-A2.
XX PD 30-JAN-2002.
XX PF 11-JUL-2001; 2001EP-00202666.
XX PR 17-JUL-2000; 2000EP-00202565.
XX PA (ALKU ) AKZO NOBEL NV.
XX PI Kusters JG, Cattoli G;
XX WPI: 2002-124384/17.
XX DR N-PSDB; ADJ58243.
XX PT Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
PT the diagnosis of Helicobacter felis infections and in the preparation of
PT vaccines.
XX PS Disclosure; SEQ ID NO 9; 76pp; English.
XX CC The present invention relates to a novel Helicobacter felis urease X and
CC Y subunit polypeptides and immunogenic fragments. The polypeptides are
CC used in the manufacture of vaccines against Helicobacter felis infections
CC and in diagnostic tests to detect antibodies against Helicobacter felis.
CC Helicobacter felis is difficult to grow so it is more convenient to use
CC the expression products of the genes encoding the urease X and Y subunits
CC in the manufacture of vaccines. The present sequence represents an
CC ureaseXY subunit of the invention.
XX
XX SQ Sequence 496 AA;
Query Match 86.3%; Score 2588; DB 5; Length 496;
Best Local Similarity 99.4%; Pred. No. 2.8e-217;
```

```
Matches 489; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 MKMKQKEYVNTYGTGDKVRLGDTDLWAEVEHDYTTYTGELKFGAGKTIREGMQSNP 60
Db 1 MKMKQKEYVNTYGTGDKVRLGDTDLWAEVEHDYTTYTGELKFGAGKTIREGMQSNP 60
QY 61 DENTLDLVTNAMIIDYTGIIKADIGIKNGKIHGIGKAGNKMDQGVSPHVVGVGTAL 120
Db 61 DENTLDLVTNAMIIDYTGIIKADIGIKNGKIHGIGKAGNKMDQGVSPHVVGVGTAL 120
QY 121 AGEGLMIITAGGIDSHTHFLSPQFPPTALANGVITTFMGGGTGPVDGTNATTITPGKWNLHR 180
Db 121 AGEGLMIITAGGIDSHTHFLSPQFPPTALANGVITTFMGGGTGPVDGTNATTITPGKWNLHR 180
QY 181 MLRAAEYSNMNVGFLGKGNSSKKQLVEQVGAIGFKLHEDWGTTPSAIDHCLSVADEY 240
Db 181 MLRAAEYSNMNVGFLGKGNSSKKQLVEQVGAIGFKLHEDWGTTPSAIDHCLSVADEY 240
QY 241 DVQVCIHDTVNEAGVDDTLNAMNGRAIHAYHIEGAGGHSPPDVITMAGELNILPSSTT 300
Db 241 DVQVCIHDTVNEAGVDDTLNAMNGRAIHAYHIEGAGGHSPPDVITMAGELNILPSSTT 300
QY 301 PTPIPTINTVAEHLDMLTCHLDKRIREDLOFSQSRIRPGSIAAEDVLHDMGVIAWTSS 360
Db 301 PTPIPTINTVAEHLDMLTCHLDKRIREDLOFSQSRIRPGSIAAEDVLHDMGVIAWTSS 360
QY 361 DSQAMGRAGEVIPRTWQTADKNKCFGLPEQKNDNFRIKRYISKYTTINPALTHGVSE 420
Db 361 DSQAMGRAGEVIPRTWQTADKNKCFGLPEQKNDNFRIKRYISKYTTINPALTHGVSE 420
QY 421 YIGSVEEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPQPVYIREMFGH 480
Db 421 YIGSVEEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPQPVYIREMFGH 480
QY 481 HGKAKFDTSTF 492
Db 481 HGKAKFDTSTF 492

RESULT 6
ADQ37851
ID ADQ37851 standard; protein; 569 AA.
XX AC ADQ37851;
XX DT 07-OCT-2004 (first entry)
XX DE H. bizzoeronii ureB polypeptide.
XX KW Urease; urease gene cluster; urease structural gene;
KW urease accessory gene; ureB; Helicobacter bizzoeronii infection;
KW antibacterial; enzyme.
XX OS Helicobacter bizzoeronii.
XX PN US2004142343-A1.
XX PD 22-JUL-2004.
XX PF 12-AUG-2003; 2003US-00639273.
XX PR 16-AUG-2002; 2002US-0404337P.
XX PA (CHAN/) CHANG Y.
XX PA (SIMP/) SIMPSON K W.
XX PA (ZHUJ/) ZHU J.
XX PI Chang Y, Simpson KW, Zhu J;
XX WPI: 2004-533502/51.
XX DR N-PSDB; ADQ37850.
XX PT Novel isolated nucleic acid molecule having urease gene cluster, and
```

conferring on *Helicobacter bizzozeronii* ability to produce urease, useful as vaccine for preventing disease in mammals infected by *H. bizzozeronii*.

Claim 24; SEQ ID NO 5; 40pp; English.

The invention relates to an isolated nucleic acid molecule conferring on *Helicobacter bizozeronii* an ability to produce urease, where the nucleic acid molecule is a urease gene cluster comprising at least one urease structural gene and at least one urease accessory gene. The nucleic acid molecule is chosen from ureA, ureB, ureE, ureF, ureG, ureH and ureI. The invention also relates to an isolated protein encoded by the nucleic acid, a vaccine for preventing onset of disease in mammals infected by *H. bizozeronii* comprising a nucleic acid and a carrier, and an isolated antibody or its binding portion raised against the nucleic acid. The nucleic acids, proteins and antibodies are useful for vaccinating mammals against onset of disease caused by infection of *H. bizozeronii*, which involves administering the sequences. The sequences are useful for detecting *H. bizozeronii* in a sample of tissue or body fluids which involves providing a nucleic acid as an antigen, providing an antibody, or providing a nucleotide sequence as a probe in a nucleic acid hybridisation assay, contacting the sample with the antigen or the probe, and detecting any reaction which indicates that *H. bizozeronii* is present in the sample. This sequence represents the *H. bizozeronii* ureB polypeptide.

Sequence 569 AA;

| | | | | |
|----------------------|--------|---------------------|-----------|-------------|
| Very Match | 77.6%; | Score 2326; | DB 8; | Length 569; |
| 1st Local Similarity | 74.1%; | Pred. No. 3.1e-194; | | |
| Conservative | 75; | Mismatches 72; | Indels 0; | Gaps 0; |

[illegible]

RESULT 7

AAR04580
ID AAR04580 standard; protein; 569 AA.

AC AAR04580;

DT 24-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 21-SEP-1990 (first entry)

XX

DE Part of protei
yy

XX KW release: probe

XX
XV
crease; probe.

OS Helicobacter pylori.

XX

PN WO9004030-A.

XX

PD 19-APR-1990.

XX
BE 06-OCT-1999. 88EB-00013135

PF
XX
06-OCT-1988; 88FR-00013133.

06-OCT-1988

XX
XX

PA (INSP) INST PAST

PA (INRM) INSERM INST NAT SANTE &

XX

PI Labigne A;

XX
DB
WDT: 1990-147914

DR WPI; 1990-147844/19.
DR N-PSDB: AAC04329.

DK N-FSDB; AAQ04323.
XX

PT New nucleotide sequences encoding Campylobacter pylori-ureaseents - and

derived vectors, transformants, protein, antibodies and probes, useful in

PT diagnosis, treatment and prevention of infections.

XXI

PS Claim 12; Page 34; 47pp; French.

XX
CC The protein can be used for the prodn of antibodies and to prepare

CC The protein can be used for the prodn. of antibodies and to prepare vaccines for the prevention /treatment of *C. pylori* infections. See also

CC vaccines for the prevention / treatment of C. pylori infections: see also
CC AAR04578-81. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 2

-OCT-2003 to standardise OS field)

XX

SQ Sequence 569 AA;

S
L

J
I
!
K

C
P
P

E
O
S
O

G

C
V
U
B

A
-
S
C

C

| Query Match | Score 2297; | DB 2; | Length 569; |
|-------------|-------------|-------|-------------|
| 76.6% | | | |

Db 123 GGLIVTAGGIDTHIFTSPOQIPAFASGVVTMIGGGTGPADGNTATTITPGRNKLKM 182
QY 182 LRAAEYSNNVFLGKGNSSKKQLVEQVAGAIKFKLHEDWGTTPSAIDHCLSVADYD 241
Db 183 LRAAEYSNNLGLAKGNASNDASLADQIEAGAIKFKIHEDWGTTPSAINHALDVADKYD 242
QY 242 VQVCITHTDVNEAGYVDDTLNMGRAIHAYHIEGAGGHSPPDVTMAGELNLPSSSTTP 301
Db 243 VQVAIHTDTLNEAGCVEDTMAAIAGRTMHTFTEGAGGCHAPDIIKVAGEHNILPASTNP 302
QY 302 TTPYTINTVAEHLDMTCHHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAMTSSD 361
Db 303 TTPFTVTNEAEHMDMLVMCHLDKSIKEDVQFADSRIRPQTAAEDTLHDMGIFSISSD 362
QY 362 SQAMGRAGEVIPRTWOTADKNKKEFKLPEDGKNDNFRIKRYISKYTINPALTHGVSEY 421
Db 363 SQAMGRGEVITRTWOTADKNKKEFKLKEEGKNDNFRIKRYLSKYTINPAIHGISEY 422
QY 422 IGSVEEGKIADLVVWNPAPFFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVVYREMFHGH 481
Db 423 VGSVEVGKVDLVLSWPAFFGVKPNMIKGGFIALSQMGDANAS IPTPQPVVYREMFPAH 482
QY 482 GKAKEDTSITFVSKVAYENGVEKGLERQVLVKNCRNITKDFKFNDKTAKITVDPKT 541
Db 483 GKAKYDANITFVSOAAVDKGIKEELGLERQVLVKNCRNITKDMQFNDTTAHIEVNPET 542
QY 542 FEVFDGKLTCKTSKPTSOVPLAORYTFF 568
Db 543 YHVFVDGKEVTSKPNKVSQAQFSIF 569

RESULT 11
ADU05536
ID ADU05536 standard; protein; 569 AA.
XX
AC ADU05536;
XX
DT 27-JAN-2005 (first entry)
XX
DE H. pylori antigenic protein HP0072.
XX
KW antibacterial; antigenic; H. pylori infection.
XX
OS Helicobacter pylori.
XX
FH Key Location/Qualifiers
FT Region 533..567
FT /note= "Immunogenic region. Region specifically claimed
FT in claim 14"
XX
PN WO2004094467-A2.
XX
PD 04-NOV-2004.
XX
PF 22-APR-2004; 2004WO-BP004255.
XX
PR 22-APR-2003; 2003EP-00450097.
XX
PA (INTE-) INTERCELL AG.
XX
PI Meinke A, Min Bui D, Nagy E, Henics T;
XX
DR WPI; 2004-775908/76.
DR N-PGDB; ADU05358.
XX
PT New hyperimmune serum reactive antigens from Helicobacter pylori, and
PT encoding nucleic acid molecules, useful for diagnosing, preventing or
PT treating H. Pylori infections.
XX
PS Claim 14; SEQ ID NO 184; 176pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule encoding a
CC hyperimmune serum reactive antigen or its fragment. The composition

CC (including the nucleic acid molecule, hyperimmune serum-reactive antigen
CC or antibody) is useful for manufacturing a medicament or pharmaceutical
CC preparation (e.g. a vaccine) for treating or preventing H. pylori
CC infections. The antigen or its fragment may also be used for isolating,
CC purifying and/or identifying an interaction partner of the hyperimmune
CC serum reactive antigen or fragment; for generating a peptide binding to
CC the hyperimmune serum reactive antigen or fragment, where the peptide is
CC selected from aptamers and spiegelmers; for manufacturing a functional nucleic acid
CC ribonucleic acid selected from ribozymes, antisense nucleic acids and
CC siRNA. The present sequence represents the amino acid sequence of an H.
CC pylori antigenic protein.
XX
SQ Sequence 569 AA;

Query Match 76.6%; Score 2297; DB 8; Length 569;
Best Local Similarity 73.4%; Pred. No. 1.1e-191; Mismatches 75; Indels 0; Gaps 0;
Matches 416; Conservative 76; Mismatches 75; Indels 0; Gaps 0;
QY 2 KMKKQYVNTYGTGDKVRLGDTDLWAEVDEHDYTYGEELKFGAGKTIREGMSGNSPD 61
Db 3 KISRKEYSVMYGPTTGDKVRLGDTDLIAVEHDYTYGEELKFGGKTLREGMSGNSPN 62
QY 62 ENTLDLVITNAMIIDYTGIIYKADIGIKNGKIHGIGKAGNKMQDGVSPHVVGVCTEALA 121
Db 63 KEELDLIITNALIVDYTYGIYKADIGIKNGKIHGIGKAGNKMQDGVKNLNVGVPATEALA 122
QY 122 GEGMIITAGGIDSHTHFLSPQOPTALANGVTMMFGGTPVDCGTNATTITPGRNKLHM 181
Db 123 GEGLVITAGGIDTHIFISPPQIPTAFASGVVTMIGGGTGPADGNTATTITPGRNKLKM 182
QY 182 LRAAEYSNNVFLGKGNSSKKQLVEQVAGAIKFKLHEDWGTTPSAIDHCLSVADYD 241
Db 183 LRAAEYSNNLGLAKGNASNDASLADQIEAGAIKFKIHEDWGTTPSAINHALDVADKYD 242
QY 242 VQVCITHTDVNEAGYVDDTLNMGRAIHAYHIEGAGGHSPPDVTMAGELNLPSSSTTP 301
Db 243 VQVAIHTDTLNEAGCVEDTMAAIAGRTMHTFTEGAGGCHAPDIIKVAGEHNILPASTNP 302
QY 302 TTPYTINTVAEHLDMTCHHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAMTSSD 361
Db 303 TTPFTVTNEAEHMDMLVMCHLDKSIKEDVQFADSRIRPQTAAEDTLHDMGIFSISSD 362
QY 362 SQAMGRAGEVIPRTWOTADKNKKEFKLPEDGKNDNFRIKRYISKYTINPALTHGVSEY 421
Db 363 SQAMGRGEVITRTWOTADKNKKEFKLKEEGKNDNFRIKRYLSKYTINPAIHGISEY 422
QY 422 IGSVEEGKIADLVVWNPAPFFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVVYREMFHGH 481
Db 423 VGSVEVGKVDLVLSWPAFFGVKPNMIKGGFIALSQMGDANAS IPTPQPVVYREMFPAH 482
QY 482 GKAKEDTSITFVSKVAYENGVEKGLERQVLVKNCRNITKDFKFNDKTAKITVDPKT 541
Db 483 GKAKYDANITFVSOAAVDKGIKEELGLERQVLVKNCRNITKDMQFNDTTAHIEVNPET 542
QY 542 FEVFDGKLTCKTSKPTSOVPLAORYTFF 568
Db 543 YHVFVDGKEVTSKPNKVSQAQFSIF 569

RESULT 12
ADS09177
ID ADS09177 standard; protein; 568 AA.
XX
AC ADS09177;
XX
DT 16-DEC-2004 (first entry)
XX
DE H. pylori urease beta subunit.
XX
KW Urease; beta subunit; UreB; gastric ulcer; peptic ulcer; gastritis;
KW stomach cancer; vaccine; antibody; immune reaction.
XX

QY 122 GEGMIITAGGIDSHTHFLSPQOFPPTALANGVTTMFGGGTGPVDGTTNATITTPGKWNLRHM 181
Db 123 GEGLIIVTAGGIDTHIFISPOQIPTAFASGVTTMIGGGTGPADGTTNATITTPGRRNLKWM 182
QY 182 LRAAEYSNMVGLGKGNSSKKQLVEQVEAGAIGFKLHEDGTTTSAIDHCLSVADVD 241
Db 183 LRAAEYSNMLGFLAKGNASNDASARDQTEAGAIGFKIHEDGTTTSAINHDLVDVADKYD 242
QY 242 VQVCIHDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGHSPPDVTMAGELNLPSTTP 301
Db 243 VQVAIHTDTLNEAGCVEDTMAALAGRTMTHTFTEGAGGCHAPDI IKVAGEHNILPASTNP 302
QY 302 TTPYTINTVAEHLDMLTCHHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAWTSDD 361
Db 303 TTPFTVNTAEHMDMLVCHHLDKSIKEDVQFADSRIRPQTAAEDTLHDMGIFTSITSD 362
QY 362 SOAMGRAGEVIRPTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSEY 421
Db 363 SOAMGRVGEVIRPTWQTADKNKKEFGRLKEEKDNDNFRIKRYLSKYITINPAIHGISEY 422
QY 422 IGSVEEGKIADLVVWNPAPFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVYREMFHGH 481
Db 423 VGSVEVGKVDLVLWSPAFGVPKNMIKGGFIALSQMGDANASITPTPQPVYREMFHGH 482
QY 482 GKAKFDTSITFVSKVAYENGVEKUGLERQVLVKNCRNITKDKFKNPKTAKITVDPKT 541
Db 483 GKAKYDRNITFVSQAAYDKGKEELGLERQVLVKNCRNITKDKMQFNDTTAHIEVNPET 542
QY 542 FEVFDGKLTCKTSKPTSOVPLAQRYTFF 568
Db 543 YHVFVDGKEVTSKPNKVSIAQLFSIF 569

RESULT 14
AEB55116
ID AEB55116 standard; protein; 570 AA.
AC AEB55116;
XX
XX
DT 22-SEP-2005 (first entry)
XX
DE Helicobacter pylori Urase B protein fragment SEQ ID 5.
XX
KW PHUR3; vaccine; helicobacter pylori infection; urease;
KW immune stimulation; antibacterial; immunostimulant; antiulcer;
KW cytostatic; antiinflammatory; gastrointestinal-gen.; gastritis;
KW peptic ulcer; carcinoma.
XX
OS Helicobacter pylori.
XX
PN WO200132014-A2.
XX
XX 10-MAY-2001.
XX
PF 01-NOV-2000; 2000WO-US030191.
XX
PR 01-NOV-1999; 99US-00431705.
XX
XX (ORAV-) ORAVAX INC.
PA
XX Kleanthous H, Londono-Arcila P, Freeman D;
PI
XX WPI; 2001-343379/36.
XX
DR N-PSDB; AEB55112, AEB55130.
XX
XX Inducing an immune response against Helicobacter in mammals, useful for
PT treating Helicobacter induced gastroduodenal diseases.
XX
XX Disclosure; SEQ ID NO 5; 63pp; English.
PS
XX The invention relates to inducing an immune response against Helicobacter
CC in a mammal, comprising mucosally administering to the mammal an
CC attenuated Salmonella vector containing a nucleic acid molecule encoding

CC a Helicobacter antigen, and parenterally administering to the mammal a
CC Helicobacter antigen. Also included is an attenuated Salmonella vector
CC comprising a nucleic acid molecule encoding a Helicobacter antigen. The
CC Helicobacter antigen is a urease, a urease subunit, or its immunogenic
CC fragment (encoded by the ureA and ureB genes). The mammal is at risk of
CC having but does not have Helicobacter infection or has a Helicobacter
CC infection. The attenuated Salmonella vector further comprises an hfrA or
CC nlrB promoter. The vector can be used in inducing an immune response
CC against Helicobacter in a mammal. The vector can be used to treat
CC Helicobacter infection. The vector and the method can be used to treat
CC Helicobacter induced gastroduodenal diseases, including acute, chronic or
CC atrophic gastritis, peptic ulcer diseases (e.g. gastric or duodenal
CC ulcers), and gastric carcinoma. Attenuated Salmonella vector PHUR3 was
CC constructed expressing UreA and UreB and including an E. coli AmpR gene.
CC The present sequence represents a protein or peptide encoded by the PHUR3
CC plasmid. NOTE: PHUR3 is represented by both AEB55112 and AEB55130, the
CC peptides expressed by PHUR3 are shown in the sequence listing to be split
CC between AEB55112 and AEB55130, yet Figure 4 shows all the peptides being
CC expressed by the one sequence, therefore all encoded peptides are cross-
CC referenced to both sequences.
XX
SQ Sequence 570 AA;

Query Match 76.1%; Score 2281; DB 4; Length 570;
Best Local Similarity 73.0%; Pred. No. 2.7e-190;
Matches 414; Conservative 76; Mismatches 77; Indels 0; Gaps 0;
QY 2 KMKKQEVNNTYPTGDKVRLGDTDLWAEVDEHYDTTYGEEKLFGAGKTIREMGOSNSPD 61
Db 4 KISRKQVSVYSGPTGDKVRLGDTDLWAEVDEHYDTTYGEEKLFGAGKTIREMGOSNSPPS 63
QY 62 ENTLDLVITNAMIIDYTGKADIGIKNGKIHGKAGNKMDDQGVSPHVVGVGTEALA 121
Db 64 KEELDLIITNALIVDYTGKADIGIKNGKIHGKAGNKMDDQGVSPHVVGVGTEALA 123
QY 122 GEGMIITAGGIDSHTHFLSPQOFPPTALANGVTTMFGGGTGPVDGTTNATITTPGKWNLRHM 181
Db 124 GEGLIIVTAGGIDTHIFISPOQIPTAFASGVTTMIGGGTGPADGTTNATITTPGRRNLKWM 183
QY 182 LRAAEYSNMVGLGKGNSSKKQLVEQVEAGAIGFKLHEDGTTTSAIDHCLSVADVD 241
Db 184 LRAAEYSNMLGFLAKGNASNDASARDQTEAGAIGFKIHEDGTTTSAINHDLVDVADKYD 242
QY 242 VQVCIHDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGHSPPDVTMAGELNLPSTTP 301
Db 244 VQVAIHTDTLNEAGCVEDTMAALAGRTMTHTFTEGAGGCHAPDI IKVAGEHNILPASTNP 303
QY 302 TTPYTINTVAEHLDMLTCHHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAWTSDD 361
Db 304 TTPFTVNTAEHMDMLVCHHLDKSIKEDVQFADSRIRPQTAAEDTLHDMGIFTSITSD 363
QY 362 SOAMGRAGEVIRPTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSEY 421
Db 364 SOAMGRVGEVIRPTWQTADKNKKEFGRLKEEKDNDNFRIKRYLSKYITINPAIHGISEY 423
QY 422 IGSVEEGKIADLVVWNPAPFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVYREMFHGH 481
Db 424 VGSVEVGKVDLVLWSPAFGVPKNMIKGGFIALSQMGDANASITPTPQPVYREMFHGH 483
QY 482 GKAKFDTSITFVSKVAYENGVEKUGLERQVLVKNCRNITKDKFKNPKTAKITVDPKT 541
Db 484 GKAKYDANITFVSQAAYDKGKEELGLERQVLVKNCRNITKDKMQFNDTTAHIEVNPET 543
QY 542 FEVFDGKLTCKTSKPTSOVPLAQRYTFF 568
Db 544 YHVFVDGKEVTSKPNKVSIAQLFSIF 570

RESULT 15
AAR74337
ID AAR74337 standard; protein; 569 AA.
XX
AC AAR74337;

QY 362 SQAMGRAGEVTPRTWOTADKNKKEFGKLPEDGKNDNFRIRKRYISKYTNIPALTHGVSEY 421
Db 363 SQAMGRVGEVITRTWOTADKNKKEFGRLKEBKNDNFRIRKRYISKYTNIPGIAHGIDY 422
QY 422 IGSVEEGKIADLVVNPAFFGVGPKIVIKGGMVVFSEMGDSNASVPTQPVPVYRPMFGHH 481
Db 423 VGSVEVGKYADLVLSPAFFGIKPNMIKGGFIALSQMGDANASITPTQPVPVYRPMFGHH 482
QY 482 GKAKFDTSITFVSKVAYENGVEKLGLEQVLVKNCRNITKQDFKNDKTAKITVDPKT 541
Db 483 GKAKFDNITFVSQAAYKAGIKEELGLDRAAPPVKNCRNITKQDKLFNDVTAHIDVNPET 542
QY 542 FEVFDGKLCSTKPTSQVPLAQRTPF 568
Db 543 YKVKVDGKEVTSKADELSLAQLNLF 569

Search completed: November 28, 2005, 08:14:37
Job time : 188.28 secs

XX 25-MAR-2003 (revised)
DT 02-NOV-1995 (first entry)
XX Helicobacter felis urease UreB.
KW Urease; UreA; UreB; vaccine; Helicobacter infection; UreI;
KW heat shock protein.
OS Helicobacter felis.
XX WO9514093-A1.
PN 26-MAY-1995.
XX 19-NOV-1993; 93WO-EF003259.
PF 19-NOV-1993; 93WO-EF003259.
PR (INSP) INST PASTEUR.
PA (INRM) INST NAT SANTE & RECH MEDICALE.
XX Labigne A, Suerbaum S, Ferrero R;
XX WPI; 1995-200383/26.
DR N-PSDB; AAQ90180.
XX Immunogenic composition against Helicobacter infection - also gene
PT fragment(s) and protein(s) from Helicobacter urease gene cluster and heat
PT shock protein(s).
XX Claim 10; Fig 3; 128pp; English.
XX The urease UreB protein is a component of a novel immunogenic composition
CC capable of inducing protective antibodies against Helicobacter infection.
CC The composition may include the UreA protein (AAR74336), a urease-
CC associated heat shock protein (AAR74338-39) or the UreI protein
CC (AAR74340). The composition is used to prepare a vaccine for humans or
CC animals, especially against H. pylori and H. felis. Antibodies against
CC the proteins may be used for treating Helicobacter infection, and
CC primers/probes to the DNA sequence may be used for detection of
CC Helicobacter infection. (Updated on 25-MAR-2003 to correct PN field.)
XX Sequence 569 AA;
Query Match 74.8%; Score 2244; DB 2; Length 569;
Best Local Similarity 72.5%; Pred. No. 4.7e-187;
Matches 411; Conservative 71; Mismatches 85; Indels 0; Gaps 0;
QY 2 KMKKQEVNTYGTGKDKVRLGDTDLWAEVEHDYTTYGBELKEGAGKTIREGMGQSNSPD 61
Db 3 KISRKEYSMYGTPTGDRVRLGDTDLILEVHDCTTYGEEIKFGGKTI RDGMSQINSPPS 62
QY 62 ENTDLVITNAMIIDYTGIVKADIGIKNGKIHGIGKAGNKMDQGVSPHVVGVGTEALA 121
Db 63 SYELDLVLNLIYDVTGYIKADIGIKGKAGNKMDQGVDDNLCVGPATEALA 122
QY 122 GEGMIITAGIDSHTHFLSPQPPPTALANGVTMTFGGTPVDGTNATTITPKWNLHRM 181
Db 123 AEGLIVTAGIDTHIHFIHQIPTAFASGVTTMIGGTPADGTNATTITPGRANLKS 182
QY 182 LRAAEYSNNVFLGKGNSSSKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVAD EYD 241
Db 183 LRAAEYANMLGFLAKGNVSYEPLSDQIEAGAIGFKIHEDWGTTPAAIHCLNVADEYD 242
QY 242 VQVCIHTDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGGHS PDVITMAGELNLPSTTP 301
Db 243 VQVAHTDTLNEAGCVEDLEAAGTTHFTHTEGAGGGHAPDVIKWAGEFNILPASTNP 302
QY 302 TIPTTINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIA MTSSD 361
Db 303 TIPTTKNTEAEHMDMLMVCHLDKSIKEDVQPADSRIRPQTIAAEDQLHDMGIFSI TSSD 362

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 28, 2005, 08:11:38 ; Search time 160.242 Seconds
(without alignments)
1481.056 Million cell updates/sec

Title: US-09-904-994B-3
Perfect score: 2999
Sequence: 1 MKMKQEVNTYGTGDKV.....KLCTSKPTSQVPLAQRYTF 568

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pgp:*
2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pgp:*
3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pgp:*
4: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pgp:*
5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pgp:*
6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pgp:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

| SUMMARIES | | | | | | | | | |
|------------|--------|---------|-----|--------|----------------------|----|--------------------|--|--|
| Result No. | Score | Query % | | Length | DB | ID | Description | | |
| | | Match | | | | | | | |
| 1 | 2999 | 100.0 | 568 | 3 | US-09-904-994B-3 | | Sequence 3, Appli | | |
| 2 | 2973 | 99.1 | 568 | 3 | US-09-904-994B-15 | | Sequence 15, Appl | | |
| 3 | 2960 | 98.7 | 568 | 3 | US-09-904-994B-6 | | Sequence 6, Appli | | |
| 4 | 2953 | 98.5 | 568 | 3 | US-09-904-994B-12 | | Sequence 12, Appl | | |
| 5 | 2588 | 86.3 | 496 | 3 | US-09-904-994B-9 | | Sequence 9, Appli | | |
| 6 | 2326 | 77.6 | 569 | 4 | US-10-639-273-5 | | Sequence 5, Appli | | |
| 7 | 2298.5 | 76.6 | 568 | 4 | US-10-639-273-38 | | Sequence 38, Appl | | |
| 8 | 2297 | 76.6 | 569 | 3 | US-09-895-913A-252 | | Sequence 252, App | | |
| 9 | 2297 | 76.6 | 569 | 4 | US-10-282-122A-58611 | | Sequence 58611, A | | |
| 10 | 2297 | 76.6 | 569 | 4 | US-10-335-977-8611 | | Sequence 8611, Ap | | |
| 11 | 2297 | 76.6 | 569 | 4 | US-10-639-273-40 | | Sequence 40, Appl | | |
| 12 | 2244 | 74.8 | 569 | 4 | US-10-639-273-39 | | Sequence 39, Appl | | |
| 13 | 2234 | 74.5 | 750 | 3 | US-09-402-100-2 | | Sequence 2, Appli | | |
| 14 | 2107 | 70.3 | 569 | 4 | US-10-639-273-41 | | Sequence 41, Appl | | |
| 15 | 1922 | 64.1 | 568 | 4 | US-10-282-122A-47638 | | Sequence 47638, A | | |
| 16 | 1914 | 63.8 | 566 | 4 | US-10-282-122A-69286 | | Sequence 69286, A | | |
| 17 | 1907 | 63.6 | 568 | 4 | US-10-282-122A-49766 | | Sequence 49766, A | | |
| 18 | 1897 | 63.3 | 568 | 4 | US-10-282-122A-50238 | | Sequence 50238, A | | |
| 19 | 1832 | 61.1 | 566 | 4 | US-10-282-122A-66743 | | Sequence 66743, A | | |
| 20 | 1826 | 60.9 | 572 | 4 | US-10-282-122A-58210 | | Sequence 58210, A | | |
| 21 | 1825 | 60.9 | 567 | 4 | US-10-282-122A-60007 | | Sequence 60007, A | | |
| 22 | 1818.5 | 60.6 | 779 | 4 | US-10-282-122A-69134 | | Sequence 69134, A | | |
| 23 | 1815 | 60.5 | 566 | 4 | US-10-282-122A-44964 | | Sequence 44964, A | | |
| 24 | 1809 | 60.3 | 568 | 4 | US-10-282-122A-56823 | | Sequence 56823, A | | |
| 25 | 1809 | 60.3 | 568 | 4 | US-10-282-122A-56844 | | Sequence 56844, A | | |
| 26 | 1807 | 60.3 | 837 | 4 | US-10-424-599-254635 | | Sequence 254635, A | | |
| 27 | 1797 | 59.9 | 567 | 4 | US-10-282-122A-68252 | | Sequence 68252, A | | |

| | | | | | | |
|----|--------|------|-----|---|----------------------|--------------------|
| 28 | 1789.5 | 59.7 | 811 | 4 | US-10-282-122A-70574 | Sequence 70574, A |
| 29 | 1788.5 | 59.6 | 573 | 4 | US-10-724-972A-6526 | Sequence 6526, Ap |
| 30 | 1784.5 | 59.5 | 571 | 4 | US-10-282-122A-44443 | Sequence 44443, A |
| 31 | 1779.5 | 59.3 | 572 | 4 | US-10-282-122A-78027 | Sequence 78027, A |
| 32 | 1753.5 | 58.5 | 840 | 4 | US-10-621-833-7 | Sequence 7, Appli |
| 33 | 1753.5 | 58.5 | 840 | 5 | US-10-731-877-1 | Sequence 1, Appli |
| 34 | 1753.5 | 58.5 | 840 | 6 | US-11-046-271-1 | Sequence 1, Appli |
| 35 | 1749 | 58.3 | 843 | 4 | US-10-437-963-124057 | Sequence 124057, A |
| 36 | 1716 | 57.2 | 570 | 3 | US-09-738-626-3599 | Sequence 3599, Ap |
| 37 | 1699.5 | 56.7 | 837 | 4 | US-10-418-962-2 | Sequence 2, Appli |
| 38 | 1696.5 | 56.6 | 598 | 4 | US-10-282-122A-76849 | Sequence 76849, A |
| 39 | 1691 | 56.4 | 571 | 4 | US-10-282-122A-51072 | Sequence 51072, A |
| 40 | 1691 | 56.4 | 573 | 4 | US-10-156-761-14635 | Sequence 14635, A |
| 41 | 1610 | 53.7 | 577 | 4 | US-10-282-122A-62532 | Sequence 62532, A |
| 42 | 1610 | 53.7 | 577 | 4 | US-10-282-122A-64617 | Sequence 64617, A |
| 43 | 1462 | 48.7 | 582 | 4 | US-10-418-962-4 | Sequence 4, Appli |
| 44 | 1241.5 | 41.4 | 556 | 4 | US-10-156-761-10252 | Sequence 10252, A |
| 45 | 1129 | 37.6 | 274 | 4 | US-10-335-977-8610 | Sequence 8610, Ap |

ALIGNMENTS

RESULT 1

US-09-904-994B-3
; Sequence 3, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO NO. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degosedesequenties
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-3

Query Match 100.0%; Score 2999; DB 3; Length 568;
Best Local Similarity 100.0%; Pred. No. 1.4e-258;
Matches 568; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| Qy | 1 | MKMKQEVNTYGTGDKVRLGDTDLWAEVEHDYTTYGEEKFGAGKTIREGMQSNP | 60 |
| Db | 1 | MKMKQEVNTYGTGDKVRLGDTDLWAEVEHDYTTYGEEKFGAGKTIREGMQSNP | 60 |
| Qy | 61 | DENTLDLVIITNMIIDYTGKADIGIKNGKIHGKAGNKMDDGVSPHVVGVGTEAL | 120 |
| Db | 61 | DENTLDLVIITNMIIDYTGKADIGIKNGKIHGKAGNKMDDGVSPHVVGVGTEAL | 120 |
| Qy | 121 | AGSGMIITAGIDISHTHFLSPQQPPTALANGVTMTMFGGCTGPDVGTNATTTTPGKNLHR | 180 |
| Db | 121 | AGSGMIITAGIDISHTHFLSPQQPPTALANGVTMTMFGGCTGPDVGTNATTTTPGKNLHR | 180 |
| Qy | 181 | MLRAAEYSNMVGLFGKNSSSKKQLVEQVEAGAIGFKLHEDWGTPPSAIDHCLSVADY | 240 |
| Db | 181 | MLRAAEYSNMVGLFGKNSSSKKQLVEQVEAGAIGFKLHEDWGTPPSAIDHCLSVADY | 240 |
| Qy | 241 | DVQVCIHDTVNEAGYVDDTLNANGRAIHAYHIEGAGGHSPPDVTWAGELNIPSSYT | 300 |
| Db | 241 | DVQVCIHDTVNEAGYVDDTLNANGRAIHAYHIEGAGGHSPPDVTWAGELNIPSSYT | 300 |
| Qy | 301 | PTTPYTINTVAEHLDMLTCHHDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAITSS | 360 |
| Db | 301 | PTTPYTINTVAEHLDMLTCHHDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAITSS | 360 |
| Qy | 361 | DSQAMGRAGEVIPRTWTQADNKKKFGKLPDGDKNNDNFRIKRYISKYITINPALTGHVSE | 420 |
| Db | 361 | DSQAMGRAGEVIPRTWTQADNKKKFGKLPDGDKNNDNFRIKRYISKYITINPALTGHVSE | 420 |

Qy 421 YIGSVEEGKIADLVVWNPFAFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVVYREMFH 480
Db 421 YIGSVEEGKIADLVVWNPFAFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVVYREMFH 480
Qy 481 HGKAKFDTSIITFVSKVAVENGVEKLGLEKRLVLPVKNCRNITKKDFKNDKTAKITVDPK 540
Db 481 HGKAKFDTSIITFVSKVAVENGVEKLGLEKRLVLPVKNCRNITKKDFKNDKTAKITVDPK 540
Qy 541 TFEVFDGKCLCTSKPTSOVPLAQRITFF 568
Db 541 TFEVFDGKCLCTSKPTSOVPLAQRITFF 568

RESULT 2
US-09-904-994B-15
; Sequence 15, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-15

Query Match 99.1%; Score 2973; DB 3; Length 568;
Best Local Similarity 98.9%; Pred. No. 2.9e-256;
Matches 562; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MKMKKQEVYNTYGTPTGDKVRLGDTDLMAEVEHDYTYGEEELKFGAGKTIREGMGQSNP 60
Db 1 MKMKKQEVYNTYGTPTGDKVRLGDTDLMAEVEHDYTYGEEELKFGAGKTIREGMGQSNP 60
Qy 61 DENTLDLVTNAMIIDYTYGIYKADIGIKNGKIHGIGKAGNKMDQGVSPHVMVGVGTEAL 120
Db 61 DENTLDLVTNAMIIDYTYGIYKADIGIKNGKIHGIGKAGNKMDQGVSPHVMVGVGTEAL 120
Qy 121 AGEGLMIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGCTGPDGTNATITTPGKNLHR 180
Db 121 AGEGLMIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGCTGPDGTNATITTPGKNLHR 180
Qy 181 MLRAAEYSNMVGFGLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVADYEY 240
Db 181 MLRAAEYSNMVGFGLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVADYEY 240
Qy 241 DVQVCIHDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGGHSPDVIIMAGELNILPSSTT 300
Db 241 DVQVCIHDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGGHSPDVIIMAGELNILPSSTT 300
Qy 301 PTIPYTIINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAWTSS 360
Db 301 PTIPYTIINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAWTSS 360
Qy 361 DSQAMGRAGEVIPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYTIINPALTHGVSE 420
Db 361 DSQAMGRAGEVIPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYTIINPALTHGVSE 420
Qy 421 YIGSVEEGKIADLVVWNPFAFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVVYREMFH 480
Db 421 YIGSVEEGKIADLVVWNPFAFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVVYREMFH 480
Qy 481 HGKAKFDTSIITFVSKVAVENGVEKLGLEKRLVLPVKNCRNITKKDFKNDKTAKITVDPK 540
Db 481 HGKAKFDTSIITFVSKVAVENGVEKLGLEKRLVLPVKNCRNITKKDFKNDKTAKITVDPK 540
Qy 541 TFEVFDGKCLCTSKPTSOVPLAQRITFF 568
Db 541 TFEVFDGKCLCTSKPTSOVPLAQRITFF 568

Db 541 TFEVFDGKCLCTSKPTSEVPLAQRITFF 568

RESULT 3
US-09-904-994B-6
; Sequence 6, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-6

Query Match 98.7%; Score 2960; DB 3; Length 568;
Best Local Similarity 98.6%; Pred. No. 4.2e-255;
Matches 560; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MKMKKQEVYNTYGTPTGDKVRLGDTDLMAEVEHDYTYGEEELKFGAGKTIREGMGQSNP 60
Db 1 MKMKKQEVYNTYGTPTGDKVRLGDTDLMAEVEHDYTYGEEELKFGAGKTIREGMGQSNP 60
Qy 61 DENTLDLVTNAMIIDYTYGIYKADIGIKNGKIHGIGKAGNKMDQGVSPHVMVGVGTEAL 120
Db 61 DENTLDLVTNAMIIDYTYGIYKADIGIKNGKIHGIGKAGNKMDQGVSPHVMVGVGTEAL 120
Qy 121 AGEGLMIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGCTGPDGTNATITTPGKNLHR 180
Db 121 AGEGLMIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGCTGPDGTNATITTPGKNLHR 180
Qy 181 MLRAAEYSNMVGFGLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVADYEY 240
Db 181 MLRAAEYSNMVGFGLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVADYEY 240
Qy 241 DVQVCIHDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGGHSPDVIIMAGELNILPSSTT 300
Db 241 DVQVCIHDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGGHSPDVIIMAGELNILPSSTT 300
Qy 301 PTIPYTIINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAWTSS 360
Db 301 PTIPYTIINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAWTSS 360
Qy 361 DSQAMGRAGEVIPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYTIINPALTHGVSE 420
Db 361 DSQAMGRAGEVIPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYTIINPALTHGVSE 420
Qy 421 YIGSVEEGKIADLVVWNPFAFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVVYREMFH 480
Db 421 YIGSVEEGKIADLVVWNPFAFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVVYREMFH 480
Qy 481 HGKAKFDTSIITFVSKVAVENGVEKLGLEKRLVLPVKNCRNITKKDFKNDKTAKITVDPK 540
Db 481 HGKAKFDTSIITFVSKVAVENGVEKLGLEKRLVLPVKNCRNITKKDFKNDKTAKITVDPK 540
Qy 541 TFEVFDGKCLCTSKPTSOVPLAQRITFF 568
Db 541 TFEVFDGKCLCTSKPASEVPLAQRITFF 568

RESULT 4
US-09-904-994B-12
; Sequence 12, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine

```
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-12

Query Match      98.5%; Score 2953; DB 3; Length 568;
Best Local Similarity 98.2%; Pred. No. 1.8e-254;
Matches 558; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 MKMKQEVVNTYGPPTGDKVRLGDTDLWAEVEHDYTTYGEELKFGAGKTIREGMOGNSP 60
DB 1 MKMKQEVVNTYGPPTGDKVRLGDTDLWAEVEHDYTTYGEELKFGAGKTIREGMOGNSP 60
QY 61 DENTLDLVTNAMIIDYTGKADIGIKNGKIHGIGKAGNKMQDGVSPHVVGVGTEAL 120
DB 61 DENTLDLVTNAMIIDYTGKADIGIKNGKIHGIGKAGNKMQDGVSPHVVGVGTEAL 120
QY 121 AGEEMIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGTGPVDGTNATTITPGKWNLHR 180
DB 121 AGEEMIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGTGPVDGTNATTITPGKWNLHR 180
QY 181 MLRAAEYSNMVGFPGKGNSSKKQLVEOVEAGAIGFKLHEDWGTTPSAIDHCLSVADYEY 240
DB 181 MLRAAEYSNMVGFPGKGNSSKKQLVEOVEAGAIGFKLHEDWGTTPSAIDHCLSVADYEY 240
QY 241 DVQVCIHTDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGSHSPDVTIMAGELNILPSSTT 300
DB 241 DVQVCIHTDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGSHSPDVTIMAGELNILPSSTT 300
QY 301 PTIPTINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAEDVLHDMGVIAWTS 360
DB 301 PTIPTINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAEDVLHDMGVIAWTS 360
QY 361 DSQAMGRAGEVTPRTWQTADKNKEFGKLPEDGKNDNFRIKRYISKYTIINPALTHGVSE 420
DB 361 DSQAMGRAGEVTPRTWQTADKNKEFGKLPEDGKNDNFRIKRYISKYTIINPALTHGVSE 420
QY 421 YIGSVEEGKIADLVVWNPFAFGVKPKIVIKGMVVFSEMGDSNASVPTPQPVYVREMFGH 480
DB 421 YIGSVEEGKIADLVVWNPFAFGVKPKIVIKGMVVFSEMGDSNASVPTPQPVYVREMFGH 480
QY 481 HGKAKFDTSIITF 492
DB 481 HGKAKFDTSIITF 492

RESULT 6
US-10-639-273-5
; Sequence 5, Application US/10639273
; Publication No. US20040142343A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; APPLICANT: Simpson, Kenneth W
; APPLICANT: Zhu, Jiaguan
; TITLE OF INVENTION: HELICOBACTER BIZZOZERONII UREASE GENES AND THEIR USES
; TITLE OF INVENTION: IN DIAGNOSTIC AND TREATMENT METHODS
; FILE REFERENCE: 19603/3881
; CURRENT APPLICATION NUMBER: US/10/639,273
; CURRENT FILING DATE: 2003-08-12
; PRIOR APPLICATION NUMBER: 60/404,337
; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Helicobacter bizzozeronii
US-10-639-273-5

Query Match      77.6%; Score 2326; DB 4; Length 569;
Best Local Similarity 74.1%; Pred. No. 1.8e-198;
Matches 420; Conservative 75; Mismatches 72; Indels 0; Gaps 0;

QY 2 KMKKQEVVNTYGPPTGDKVRLGDTDLWAEVEHDYTTYGEELKFGAGKTIREGMOGNSPD 61
DB 3 KISRKEYVSMYGPPTGDKVRLGDTDLILEVHDCTTYGEEKIKFGGGKTIRDMGMAOTNSPS 62
```



```
QY 182 LRAAEYSNMVGLFKGNSSKKQLVEQVEAGAIGFKLHEDMCTTPSAIDHCLSLVADEYD 241
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 183 LRAAEYSNMVGLFKGNASNDASLADQIEAGAIGFKLHEDMCTTPSAINHDLVDADKYD 242
QY 242 VQVCIHTDVTNBRAGYVDDTLNAMNGRAIHAYHIEGAGGHSPPVITMAGELNLPSTTP 301
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 243 VQVAIHDTLNEAGCVEDTMAAIAGRTHMTFTHTGAGGGHAPDIIKVAAGEHNILPASTNP 302
QY 302 TTPYINTVAEHLDMTCHLDRKIREDLQSQSRIKSGSIAAEDVLHDMGVIAVTSSD 361
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 303 TTPFTVNTAEHMDMLVCHLDKSIKEDVQFADSRIRPQTIAAEDTLHDMGIFSITSSD 362
QY 362 SQAMGRAGEVIPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYTTINPALTHGVSEY 421
Db 363 SQAMGRVGEVITRTWQTADKNKKEFGRLKEBKNDNFRIKRYLSKYTTINPAIHGISEY 422
QY 422 IGSVEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPQPVYVREMGHH 481
Db 423 VGSVEVGKADLVLSNPAFFGKPNMIKGGFIALSQMGDANASIPTPQPVYVREMFHH 482
QY 482 GKAKEDTSITFVSKVAYENGVEKGLERQVLPVKNCRNITKKDFKNDKTAKITVDPKT 541
Db 483 GKAKYDANITFVSQAAYDKGIEELGLERQVLPVKNCRNITKKDMQFNDTTAHIEVNPET 542
QY 542 FEVFDGKLCSTKPTSOVPLAORYTF 568
Db 543 YHVFVDGKEVTSKPANKVSLAQLFSIF 569
```

RESULT 9

```
US-10-282-122A-58611
; Sequence 58611, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
```

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 58611
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Helicobacter pylori
; US-10-282-122A-58611
```

```
Query Match 76.6%; Score 2297; DB 4; Length 569;
Best Local Similarity 73.4%; Pred. No. 7.2e-196;
Matches 416; Conservative 76; Mismatches 75; Indels 0; Gaps 0;
```

```
QY 2 KMKKQEVNVTGPTKGDVRLGDTDLMAEVEHDYTYGEEELKFGAGKTIREGMGQSNSPD 61
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3 KISRKEYVSMGPTTGDVRLGDTDLIAEVEHDYTYGEEELKFGGKTIREGMSQSNPS 62
QY 62 ENTLDLVTNAMIIDYTGIVKADIGIKNGKIHGKAGNKMDQGVSPHVMVGVGTTEALA 121
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 63 KEELDLITNALIVDYGIVKADIGIKDGKIAGIKGNGKMDQGVKNNLSVGPATEALA 122
QY 122 GEGMIITAGGIDSHTEFLSPQOFPFTALANGVTTFMFGGCTGPDGTNATTTITPGKWNLHRM 181
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 123 GEGLIVTAGGIDTHIHFIISPOIIFTFASGVTTMIGGCTGPDGTNATTTITPGRNKLWM 182
QY 182 LRAAEYSNMVGLFKGNSSKKQLVEQVEAGAIGFKLHEDMGTTPPSAIDHCLSLVADEYD 241
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 183 LRAAEYSNMVGLFKGNASNDASLADQIEAGAIGFKIHEDMGTTPPSAINHALDVADKYD 242
QY 242 VQVCIHTDVTNBRAGYVDDTLNAMNGRAIHAYHIEGAGGHSPPVITMAGELNLPSTTP 301
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 243 VQVAIHDTLNEAGCVEDTMAAIAGRTHMTFTHTGAGGGHAPDIIKVAAGEHNILPASTNP 302
QY 302 TTPYINTVAEHLDMTCHLDRKIREDLQSQSRIKSGSIAAEDVLHDMGVIAVTSSD 361
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 303 TTPFTVNTAEHMDMLVCHLDKSIKEDVQFADSRIRPQTIAAEDTLHDMGIFSITSSD 362
QY 362 SQAMGRAGEVIPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYTTINPALTHGVSEY 421
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 363 SQAMGRVGEVITRTWQTADKNKKEFGRLKEBKNDNFRIKRYLSKYTTINPAIHGISEY 422
QY 422 IGSVEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPQPVYVREMGHH 481
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 423 VGSVEVGKADLVLSNPAFFGKPNMIKGGFIALSQMGDANASIPTPQPVYVREMFHH 482
QY 482 GKAKEDTSITFVSKVAYENGVEKGLERQVLPVKNCRNITKKDFKNDKTAKITVDPKT 541
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 483 GKAKYDANITFVSQAAYDKGIEELGLERQVLPVKNCRNITKKDMQFNDTTAHIEVNPET 542
QY 542 FEVFDGKLCSTKPTSOVPLAORYTF 568
Db 543 YHVFVDGKEVTSKPANKVSLAQLFSIF 569
```

RESULT 10

```
US-10-335-977-8611
; Sequence 8611, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
```


| | | | | | | | | | | | | |
|----|-----|----------|---------------|-------|--------|-------|-------|------|-----|-------|--------|-------------------|
| QY | 481 | HGKAKFDT | SITFVSKVAYENG | KEKLG | LERQVL | LVKNC | RNITK | KDKF | KFN | DKTAK | ITVDPK | 540 |
| Db | | | | | | | | | | | | : |
| | | | | | | | | | | | | : |
| Db | 481 | RGGALART | SLTFV | SQMA | DAGIA | ERYGL | AKRI | VPVR | NC | RNVTK | ADMI | HNAWRPSISVDPE 540 |
| | | | | | | | | | | | | : |
| | | | | | | | | | | | | : |
| QY | 541 | TFEV | FVDG | KLCT | SKPTS | QVPL | AQRY | TF | | | | 568 |
| Db | | | | | | | | | | | | : |
| | | | | | | | | | | | | : |
| Db | 541 | TYDVI | ADG | QLT | CEPAT | VLP | MAQRY | FLF | | | | 568 |
| | | | | | | | | | | | | : |
| | | | | | | | | | | | | : |

Search completed: November 28, 2005, 08:24:32
Job time : 162.242 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 28, 2005, 08:14:48 ; Search time 7.15365 Seconds
(without alignments)
240.659 Million cell updates/sec

Title: US-09-904-994B-3
Perfect score: 2999
Sequence: 1 MKKKQYVNTYPTKGDV.....KLCTSKPTSQVPLAQRYTFF 568

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 17545 seqs, 3030971 residues

Total number of hits satisfying chosen parameters: 17545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA_New:*
1: /cgn2_6/prodata/1/pubpa/US10_NEW_PUB.pap:*
2: /cgn2_6/prodata/1/pubpa/US06_NEW_PUB.pap:*
3: /cgn2_6/prodata/1/pubpa/US07_NEW_PUB.pap:*
4: /cgn2_6/prodata/1/pubpa/US08_NEW_PUB.pap:*
5: /cgn2_6/prodata/1/pubpa/US09_NEW_PUB.pap:*
6: /cgn2_6/prodata/1/pubpa/PCT_NEW_PUB.pap:*
7: /cgn2_6/prodata/1/pubpa/US11_NEW_PUB.pap:*
8: /cgn2_6/prodata/1/pubpa/US60_NEW_PUB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|--------------------|
| 1 | 1789.5 | 59.7 | 571 | 1 | US-10-793-626-118 |
| 2 | 1380.5 | 46.0 | 431 | 1 | US-10-485-517-182 |
| 3 | 123.5 | 4.1 | 802 | 1 | US-10-510-386-2 |
| 4 | 113 | 3.8 | 2314 | 7 | US-11-013-759-11 |
| 5 | 105.5 | 3.5 | 801 | 1 | US-10-793-626-90 |
| 6 | 100 | 3.3 | 403 | 7 | US-11-074-176-354 |
| 7 | 100 | 3.3 | 408 | 7 | US-11-074-176-186 |
| 8 | 99.5 | 3.3 | 1992 | 7 | US-11-013-759-3 |
| 9 | 99.5 | 3.3 | 1992 | 7 | US-11-013-759-13 |
| 10 | 99.5 | 3.3 | 2047 | 7 | US-11-013-759-7 |
| 11 | 99.5 | 3.3 | 2047 | 7 | US-11-013-759-7 |
| 12 | 96 | 3.2 | 1160 | 1 | US-10-131-826A-234 |
| 13 | 94.5 | 3.2 | 480 | 1 | US-10-821-234-1465 |
| 14 | 94 | 3.1 | 2053 | 7 | US-11-013-759-9 |
| 15 | 93 | 3.1 | 1047 | 1 | US-10-510-386-200 |
| 16 | 93 | 3.1 | 1437 | 7 | US-11-074-176-96 |
| 17 | 90.5 | 3.0 | 585 | 1 | US-10-793-626-2124 |
| 18 | 90 | 3.0 | 547 | 7 | US-11-082-389-182 |
| 19 | 89.5 | 3.0 | 491 | 1 | US-10-793-626-2808 |
| 20 | 89 | 3.0 | 770 | 1 | US-10-821-234-1269 |
| 21 | 88 | 2.9 | 2323 | 1 | US-10-793-626-760 |
| 22 | 86 | 2.9 | 469 | 1 | US-10-793-626-2204 |
| 23 | 85.5 | 2.9 | 1267 | 7 | US-11-109-156-35 |
| 24 | 84.5 | 2.8 | 530 | 1 | US-10-858-730-67 |
| 25 | 84.5 | 2.8 | 3717 | 1 | US-10-821-234-1076 |

ALIGNMENTS

RESULT 1

US-10-793-626-118
; Sequence 118 Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10793.626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 118
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-118

| | | | | |
|-----------------------|-------|---|-----------------|-------------|
| Query Match | 59.7% | Score 1789.5; | DB 1; | Length 571; |
| Best Local Similarity | 58.2% | Pred. No. 7.2e-135; | Mismatches 141; | Indels 5; |
| Matches | 332; | Conservative 92; | Indels 5; | Gaps 2; |
| Qy | 2 | MMKKQYVNTYPTKGDVRLGDTDLMAEVEHDYTYGEEKLFGAGKTIREGMGQS--- | N 58 | |
| Db | 4 | KMTQSYTSLSYPTGPTGDSVRLGDTNLFQAQVEKDYANTYGDATFGGKKSIRDMGACNPNT | 63 | |
| Qy | 59 | SPENTLDLVITNAMIIDYGYIYKADIGIKNGKHGICAGKNDQDGVSPHMYVGCTE | 118 | |
| Db | 64 | RDDKNVADLVLTNALIIDYKIVKADIGIKNGYIFKIGKAGNPIDMNDV--IIIGATTD | 121 | |
| Qy | 119 | ALAGEMLITAGGDSHTFLSPQFPPTALANGVTMTFGGTPGVDGTNATTITPKGNL | 178 | |
| Db | 122 | IIAAGKIVTAGGIDTTHVHFINPEQAEVALSGGITHHGGTGASEGAKATTITFGPHI | 181 | |
| Qy | 179 | HRMLRAAEYSMNVGLGKGNSSKKQVEQAGEAIGFKLHEDWGTTFPSAIDHCLSVAD | 238 | |
| Db | 182 | HRMLRAAEEMPINVGFTKGQAVNHTALIEQHAGAIGLVKVEDWGAFTPSALSHALDVAD | 241 | |
| Qy | 239 | EYDVQVCITDVTNAGYVDDTLNAMNGRAIHAYHIEGAGGCHSPDVTIMAGELNLPSS | 298 | |
| Db | 242 | EFDVQVALHADTLNEAGFMEDTMAAVKDRVLHMYHTEGAGGGHAPDLIKSAAYSNI | 301 | |
| Qy | 299 | TTPTIPTINTVAEHLDMLTCHHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGIAMT | 358 | |

Db 602 SYGTFL--KDKGKQTKQAPTIENLSHRKAY---QLEYSFKGTGITVS--GTERVVVPA 654
QY 469 PQVYVYREMFHGHGKAFDTSITFSKVAYENGVEKELGL-----ERQVLPVKN 518
Db 655 NQT-----GKAATVTV-----NSAKTKAGTYEYVIREGRKVAEIPLL 695
QY 519 RNITKDKDFKNDYAKITVDP-----KTFEYFVDG 548
Db 696 LIVKEPDY---PRVTSVTFPGAKQAGAYTIEAYLP 728

RESULT 4
US-11-013-759-11
; Sequence 11, Application US/11013759
; Publication No. US20050249747A1
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Sasaki, Ken
; APPLICANT: Yang, Yan Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE
; FILE REFERENCE: 1038-921MIS:jb
; CURRENT APPLICATION NUMBER: US/11/013,759
; CURRENT FILING DATE: 2004-12-16
; PRIOR APPLICATION NUMBER: US/09/361,619
; PRIOR FILING DATE: 1999-07-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 11
; LENGTH: 2314
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-11-013-759-11

Query Match 3.8%; Score 113; DB 7; Length 2314;
Best Local Similarity 20.0%; Pred. No. 0.47;
Matches 128; Conservative 68; Mismatches 201; Indels 244; Gaps 31;
QY 8 YNVTY-----GPTGDKVRGLGDTLWAEVHDYTYGE--ELKFGAGKTIREGMCQSN 58
Db 779 FVSTYNTVDFIDGNATTAFTYDTEQTQSKVTYDNNVDEKTIETLTDGNGKTNIGV--- 834
QY 59 SPENTLDIVITNAMIIDYGIYKADIGIKNGIKHIGKAGN-----KMDQGVSP--- 109
Db 835 -----KTTLLTTNA-----NGKATFSTTDNALVNAKDIAENLNTLAK 874
QY 110 --HMVVGVTGTEAL-----AGEGMIITAGGIDSHTHPLSPQFPPTALANGVTTFMG 158
Db 875 EIHITKGTADTALQTPKVKDGATDDEITVG-----KDGTONKTV--- 916
QY 159 GTPGVDTNATTTTPGK-----WNLHRLRAAEYSNMVGFGLKGNSSSKQLVEQVE 211
Db 917 NTLKLGKGLTVATNKDGTVTFTGINTQSLKAGDSTTLNKGGLSKNPASN-----BQIQ 972
QY 212 AGAIGPKLHE--DNGTTPSAIDHCLSVADEYDVQVCIHTDTVNEAGYVDDTLNAMNGRAIH 270
Db 973 VGADGVKFAKVDKNSSTGIDGTSRI-----TKQIQGFT----- 1006
QY 271 AVHIEGAGGHSPODVTMAGELNPSSTPTTPTINTVVAEHLDMLTCHLDKRIRED 330
Db 1007 -----GANG-----SLDTTKPHLTK-----DKLKVE 1028
QY 331 LQFSQSRIRPGSIAAEVDLHDMGVIAWMTSSDSQAMGRAGEV-----IPRTWQTADKNK 383
Db 1029 VEITNTGINAGGKIITNI--QSGDITQNSNDVATGGRVYDLKTELESKINSAAKTAQNSL 1086
QY 384 KEFGKLPEDGKDNNDNFRKRYISKYITINPALTHGVSEYIGSVBEGKIADLVWNPAPFGV 443
Db 1087 HEFVSVADEQC---NHFTVSNPYSYD-----TSKTSDVITFAGENGITTKVNGGVVRGI 1138
QY 444 -----KPKIVIKG---GMVVFSEMGDSN-----ASVPTPQPVYVREMGH---H 481

Db 1139 DOTKGLTTPKLTGVNNGKGIVIDSKOQNTITGLSNTLANVT-----NDGAGHALSQ 1191
QY 482 GKAKFDT-----SITFVS-----KVAYENGVEKELGLERQVLPV 515
Db 1192 GLAN-DTDKTRAASIGDVLNAGFNQLQNGEAVDFVSTYDTVTDFIDG----- 1236
QY 516 KNCENITKDKPFKND--KTAK-----ITVDPKTFEYFVDGKL 550
Db 1237 ----NATTAKVTYDDTSKTSKVYDVNVNDNKTIETVTSKKL 1273

RESULT 5
US-10-793-626-90
; Sequence 90, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 90
; LENGTH: 801
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-90

Query Match 3.5%; Score 105.5; DB 1; Length 801;
Best Local Similarity 19.5%; Pred. No. 0.39;
Matches 128; Conservative 75; Mismatches 244; Indels 211; Gaps 29;
QY 22 LGDTDLWAEVHDYTYGEEELKFGAGKTIREGMCQSNPDENTLDLVTNAMIIDYTYGIY 81
Db 214 IGDKDV-----EALKLADGRVIE-----TDLVMAVGIRPYTEVA 248
QY 82 KADTIGI-----KNGKIHGIGKAGNKD-----MDQGVSPHVMVVG 116
Db 249 K-DSGLDVNRGIVNDYMQTSDSHIYAVGECAEHDGKYGVLVAPLYEQGKVLADYLTGKE 307
QY 117 TEALAGSGM-----IITAGGI--DSHTH-----FLSPQQPPTALA 149
Db 308 TKGYKGSTTFTSLKVSQCDLYSAGQIVDEDVHGVEIFNSVDNIYKVVLSQGVVGAVL 367
QY 150 NGVTTFMGGTGTPVDGNTNATTITPKKNLHRLRAAEYSNMVGFGLKGNSSSKQLVEQ 209
Db 368 YGDTD-----DGSRF-----YNNMKKHETLEDYTL-VSLHKGDEDEAGTSIADM 410
QY 210 VEAGAIGPKLHEDMGTTPSAI--DHCLSVADEYDVQVCIHTDTVNEAGYVDDTLNAMNGRA 268
Db 411 SDEFTICGCGVDKGTIVNAITSKGLTSVDE-----VTKATKAGNSCGKKGQIGELLQYT 466
QY 269 IHAYHIEGAGGHSPP-----DVTMAGELNIPSSST-----PTIPY 305
Db 467 LGDDFIAAKPTGICPCTDLTRDQIVTQIRAKNLKSSKEVRHVLDKDKDQCPKCPAINY 526
QY 306 TINTV--AEHLDMLTCHLDKR-----IREDLQFSQSRIRPGSIAAEVDLHDMGV 354
Db 527 YLNMVYFPEHRD-----EKDSRFANERYHANIQNDGTFSVIPQMRGGVTDADQJIRLGE 580
QY 355 IA-----WTSSDSQAMGRAG---EVIPTWQTAD-KNKKEFGKLPEDGKDNNDNFRKIR 403
Db 581 VAKKYNVPLVKVTSQVRVGLYGLKKEELPQWKLGMRSASAYGKTKTSKSCVCKEPCR 640
QY 404 YISKYITINPAL-----THGVSEYIGSVBEGKIADLVWNPAPFGVVKPKTV 448
Db 641 FGQTYTTLRLGIRLEKTFEYIDTPHKFKNGVSGCPRSCVSEGVKD-----FGV---IS 689

```

RESULT 7
US-11-074-176-186
; Sequence 186, Application US/11074176
; Publication No. US20050250135A1
; GENERAL INFORMATION:
; APPLICANT: Klaenhammer, Todd R.
; APPLICANT: Russell, William M.
; APPLICANT: Altermann, Eric
; APPLICANT: McAuliffe, Olivia
; APPLICANT: Peril, Andrea Azcarate
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding
; TITLE OF INVENTION: Stress-Related Proteins and Uses Therefore
; FILE REFERENCE: 5051-694

```

| | | | | |
|-----------------------|------------------|-----------------|-------------|--------------|
| Query Match | 3.3% | Score 99.5; | DB 7; | Length 1992; |
| Best Local Similarity | 21.8% | Pred. No. 4.4; | | |
| Matches 114; | Conservative 49; | Mismatches 160; | Indels 201; | Gaps 30; |

| | | | | |
|----|-----|------------------------|--|------|
| Qy | 54 | MGQSNPDPE-NTLDLVITNAMI | DYTGVIKADIGIKNGKI-HGIG-----KAGNKMDQG | 106 |
| | | : : | : | |
| Db | 921 | VGQKXANNQVNTLT | LGKENG LNI-----KTD--KNGTVFGINTTSLGKAGKSTLNDG | 971 |
| | | : | : | |
| Qy | 107 | VSPHMVGVGCTEALAGEBGM | I-ITAGGIDSHTHFLSPQQPPTALANGVTTFMFGGGTGPDVG | 165 |
| | | : | : | |
| Db | 972 | -----GLSINKPTSGEQT | OVAGDGV-----KFAKVNNGV-----CAG-IDG | 1008 |
| | | : | : | |

QY 166 TNATTITPKGNLHRLRAAEYSYMNVLGKGNSSKK-----OLVQVEA 212
DB 1009 T--TRITRDE-----IGFTGNGSLDKSPHLSKDGINAGGKKTNIQS 1050
QY 213 GAIGFKLHEDWGTTPSAIDHCLSLVADEYDVQVCIHTDTVNEAGYVDDTLNMGRAIHAY 272
DB 1051 GEIAQNSHD--AVTGKI-----YDLK-----TELENKISSTAKT--AQN--SLHEF 1091
QY 273 HIEGAGGGH-----SPDVIITMAGELNLPSSSTTPTIPTYIN--TVAEHLMDL 317
DB 1092 SVADQGNFTVSNPNYSSYDTSKTSVDITFAGENGITTKVKNKGVRVGIDQTKGLTTPKL 1151
QY 318 MTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAMTSSDSQAMGRAGEVIPRTWQ 377
DB 1152 TVGNNGKGLVIDSQNGQNTITGLSNTLANVNDKGSVRTTEQ-----GNII----- 1198
QY 378 TADKNKKEFGKLPEDGKNDNFRIKRYISKYITNPALTHGVSEYIGSVBEGKIADLVVWN 437
DB 1199 -----KDEKTR-----AASIVD----- 1211
QY 438 PAFFGVKPKIVIKGVMVFSEMGDSNA--SVPTPOPVYVREMFGHHGKAKF---DTSITF 492
DB 1212 -----VLSAG---FNLOGGEAVDFVSTYDVTNPFAD--GNATTAKVYDDTSKT- 1255
QY 493 VSKVAYENGVEKGLGLERQVLPVKNCRNITKKDFKNDKTAKIT 536
DB 1256 -SKVYVDNVDD-----TTIEVKDKKLGKVTITLT 1284
RESULT 9
US-11-013-759-13
; Sequence 13, Application US/11013759
; Publication No. US20050249747A1
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Sasaki, Ken
; APPLICANT: Yang, Yan Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE
; FILE REFERENCE: 1038-921MIS:jb
; CURRENT APPLICATION NUMBER: US/11/013,759
; PRIOR FILING DATE: 2004-12-16
; PRIOR APPLICATION NUMBER: US/09/361,619
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 1992
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-11-013-759-13
Query Match 3.3%; Score 99.5; DB 7; Length 1992;
Best Local Similarity 21.8%; Pred. No. 4.4;
Matches 114; Conservative 49; Mismatches 160; Indels 201; Gaps 30;
QY 54 MGQSNPDE-NTLDLVTNAMIIDYTGIVKADIGIKNGKI-HGIG-----KAGNKDMQDG 106
DB 921 VQQKNNQVNTLTGKGNLNI-----KTD---KNGVTFTGINTTSLGKAGKSTLNDG 971
QY 107 VSPHVMVGVGTTEALAGEGMI-ITAGGIDSHTHFLSPQQPPTALANGVTTMFGGCTGPDVG 165
DB 972 -----GLSIKNPTGSEQIQVGADGV-----KFAKVNNGVV-----GAG-IDG 1008
QY 166 TNATTITPKGNLHRLRAAEYSYMNVLGKGNSSKK-----OLVQVEA 212
DB 1009 T--TRITRDE-----IGFTGNGSLDKSPHLSKDGINAGGKKTNIQS 1050
QY 213 GAIGFKLHEDWGTTPSAIDHCLSLVADEYDVQVCIHTDTVNEAGYVDDTLNMGRAIHAY 272
DB 1051 GEIAQNSHD--AVTGKI-----YDLK-----TELENKISSTAKT--AQN--SLHEF 1091

QY 273 HIEGAGGGH-----SPDVIITMAGELNLPSSSTTPTIPTYIN--TVAEHLMDL 317
DB 1092 SVADQGNFTVSNPNYSSYDTSKTSVDITFAGENGITTKVKNKGVRVGIDQTKGLTTPKL 1151
QY 318 MTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAMTSSDSQAMGRAGEVIPRTWQ 377
DB 1152 TVGNNGKGLVIDSQNGQNTITGLSNTLANVNDKGSVRTTEQ-----GNII----- 1198
QY 378 TADKNKKEFGKLPEDGKNDNFRIKRYISKYITNPALTHGVSEYIGSVBEGKIADLVVWN 437
DB 1199 -----KDEKTR-----AASIVD----- 1211
QY 438 PAFFGVKPKIVIKGVMVFSEMGDSNA--SVPTPOPVYVREMFGHHGKAKF---DTSITF 492
DB 1212 -----VLSAG---FNLOGGEAVDFVSTYDVTNPFAD--GNATTAKVYDDTSKT- 1255
QY 493 VSKVAYENGVEKGLGLERQVLPVKNCRNITKKDFKNDKTAKIT 536
DB 1256 -SKVYVDNVDD-----TTIEVKDKKLGKVTITLT 1284
RESULT 10
US-11-013-759-4
; Sequence 4, Application US/11013759
; Publication No. US20050249747A1
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Sasaki, Ken
; APPLICANT: Yang, Yan Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE
; FILE REFERENCE: 1038-921MIS:jb
; CURRENT APPLICATION NUMBER: US/11/013,759
; CURRENT FILING DATE: 2004-12-16
; PRIOR APPLICATION NUMBER: US/09/361,619
; PRIOR FILING DATE: 1999-07-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2047
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-11-013-759-4
Query Match 3.3%; Score 99.5; DB 7; Length 2047;
Best Local Similarity 21.8%; Pred. No. 4.6;
Matches 114; Conservative 49; Mismatches 160; Indels 201; Gaps 30;
QY 54 MGQSNPDE-NTLDLVTNAMIIDYTGIVKADIGIKNGKI-HGIG-----KAGNKDMQDG 106
DB 976 VQQKNNQVNTLTGKGNLNI-----KTD---KNGVTFTGINTTSLGKAGKSTLNDG 1026
QY 107 VSPHVMVGVGTTEALAGEGMI-ITAGGIDSHTHFLSPQQPPTALANGVTTMFGGCTGPDVG 165
DB 1027 -----GLSIKNPTGSEQIQVGADGV-----KFAKVNNGVV-----GAG-IDG 1063
QY 166 TNATTITPKGNLHRLRAAEYSYMNVLGKGNSSKK-----OLVQVEA 212
DB 1064 T--TRITRDE-----IGFTGNGSLDKSPHLSKDGINAGGKKTNIQS 1105
QY 213 GAIGFKLHEDWGTTPSAIDHCLSLVADEYDVQVCIHTDTVNEAGYVDDTLNMGRAIHAY 272
DB 1106 GEIAQNSHD--AVTGKI-----YDLK-----TELENKISSTAKT--AQN--SLHEF 1146
QY 273 HIEGAGGGH-----SPDVIITMAGELNLPSSSTTPTIPTYIN--TVAEHLMDL 317
DB 1147 SVADQGNFTVSNPNYSSYDTSKTSVDITFAGENGITTKVKNKGVRVGIDQTKGLTTPKL 1206
QY 318 MTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAMTSSDSQAMGRAGEVIPRTWQ 377
DB 1207 TVGNNGKGLVIDSQNGQNTITGLSNTLANVNDKGSVRTTEQ-----GNII----- 1253

```
QY 378 TADKNNKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSEYIGSVBEGKIADLVVWN 437
Db 1254 -----KQEDKTR-----AASIVD-----1266

QY 438 PAFFGVKPKIVIKGWMVFSEMGDSNA--SVTPQPVVYVREMGHHGKAKF---DTSITF 492
Db 1267 ---VLSAG---FNLQNGEAVDFVSYDTYTNFAD--GNATTAKVTYDDTSKT- 1310

QY 493 VSKVAYENGVEKGLGLERQVLVKNCRNITKKDFKFNKTAKIT 536
Db 1311 -SKVYDVNVDD-----TTIEVKDKLGVKTTILT 1339

RESULT 11
US-11-013-759-7
; Sequence 7, Application US/11013759
; Publication No. US20050249747A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C128
; CURRENT APPLICATION NUMBER: US/10/131,826A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; Remaining prior application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 234
; LENGTH: 1160
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-131-826A-234

Query Match 3.2%; Score 96; DB 1; Length 1160;
Best Local Similarity 18.1%; Pred. No. 3.8;
Matches 99; Conservative 81; Mismatches 162; Indels 206; Gaps 31;

QY 16 KGDVRLGDT---DLMAVEHDYTYGEEKLFGA-----GKTIREGMOQSNDPENTLDL 67
Db 66 KSDKNRIGGTIVKTYKEYKDD--SYTDEVAQPAWLGLFGLPVLQAEVG-----DV 113

QY 68 VI-----TNAMIDVTGI-YKADIGIKNGKIHGKAGKMDQDGVSPHVMVVGTEA 119
Db 114 ILIHLNKFATRPYTIHPHGVFYEKD---SEGSLYPDGSSGLKADDSVPP-----160

QY 120 LAGEGMIITAGGIDSHTH-FLSPQPFPTALANGVTITWFGGCTGPDVDTNATTITPKWNL 178
```

```
QY 493 VSKVAYENGVEKGLGLERQVLVKNCRNITKKDFKFNKTAKIT 536
Db 1311 -SKVYDVNVDD-----TTIEVKDKLGVKTTILT 1339

RESULT 12
US-10-131-826A-234
; Sequence 234, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C128
; CURRENT APPLICATION NUMBER: US/10/131,826A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; Remaining prior application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 234
; LENGTH: 1160
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-131-826A-234

Query Match 3.2%; Score 96; DB 1; Length 1160;
Best Local Similarity 18.1%; Pred. No. 3.8;
Matches 99; Conservative 81; Mismatches 162; Indels 206; Gaps 31;

QY 16 KGDVRLGDT---DLMAVEHDYTYGEEKLFGA-----GKTIREGMOQSNDPENTLDL 67
Db 66 KSDKNRIGGTIVKTYKEYKDD--SYTDEVAQPAWLGLFGLPVLQAEVG-----DV 113

QY 68 VI-----TNAMIDVTGI-YKADIGIKNGKIHGKAGKMDQDGVSPHVMVVGTEA 119
Db 114 ILIHLNKFATRPYTIHPHGVFYEKD---SEGSLYPDGSSGLKADDSVPP-----160

QY 120 LAGEGMIITAGGIDSHTH-FLSPQPFPTALANGVTITWFGGCTGPDVDTNATTITPKWNL 178
```

Db 161 -----GG--SHLYNWTIPE-----CHAPTDADPACL-----TWIY 189
QY 179 HRMLRAAEYSNMVGFILG-----KGNSSSKQ-----LVEQVEAGAIGFKLHE 221
Db 189 HSHVDAPRDIA--TGLIGPLITCKRGALDGNSPQROVDHDFLLFSVVDENLSWHLNE 246
QY 222 DWGT-----TPSAIDHCLSVADYDVQVCIHTDTVNEAGYVDDTLNAMN--GRAIHAYHIEG 276
Db 247 NIATYCDSPASVD-----KEDETFOESNRMHAIN--GFVFGNLPCLNMCAKRVAMHLFG 299
QY 277 AGG-----CHSPDVTMAGEINILPSS--TPTPTTYTINTVAEHLDM 317
Db 300 MGNEIDVHTAFFHGQMLTTRGHHTDV-----ANIFPATFTVAMVWPEPGT-----WL 347
QY 318 MTCHILDKRREDLQ-----FSQSRIRPGSIAAEEDVLHDMGVIAMTSSDS 362
Db 348 ISC-QVNSHFRDGMQALYKVKSCSMAPPVOLLTKVKRQYFIEAHEIQWDYGPWCHDSTG 406
QY 363 QAMGRAGEVIPRTWQTADKNKKEFG-----KLPEDKGDNDFRIKRYISKYTIINPALTHGV 418
Db 407 KNLREPGSISDKFFQ--KSSSRIGGTYWKVRYEAFQDETFOEKHLEE-----452
QY 419 SEVIGSVBECKIADLVVNPAPFGVKPKIVIKGMVVFSEMGD-----SNASVP---T 468
Db 453 DRHLGIL-----GPIVIRAEVGDITQVVFYFNRSQPFPSMQ 486
QY 469 PQPVVYRE 476
Db 487 PHGVFYEK 494

RESULT 13

US-10-821-234-1465
; Sequence 1465, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCES: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt SEQ_genes Version 1.0
; SEQ ID NO 1465
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1465

Query Match 3.2%; Score 94.5; DB 1; Length 480;
Best Local Similarity 19.8%; Pred. No. 1.4;
Matches 66; Conservative 48; Mismatches 95; Indels 127; Gaps 18;
QY 50 IREGMGQSNPDENTLDLVITNAMIIDYTGIVKADIGIKNGKHGIGKAGNKDMQD---- 105
Db 201 VASGSFVISED-----GLIVTNAHV--TNKRVKVELKNGAY--EAKIKDVEKADI 251
QY 106 -----GVSHPMVGVGTEALAGEMIITAGIDSHTHFLSPQFPPTALANGVTT--- 154
Db 252 ALIKIDHOGKLPVLLGRSSELPGE-FVVAIGS-----PFSLQNTVTTGIV 297
QY 155 -----MFGGSGTGP-----VDGTNATITTPG-----K 175
Db 298 STTQGGKELGRNSDMDYIQTDAIINYGNSSGGLVNLGDGEVIGINTLKTAGISFAIPS 357
QY 176 WNLHMLRAAEYSNMVGFILGKGNSSKKQLVQVEAGAIGFKL-----HE 221

Db 358 DKIKKFLTESHDRO-----AKGKAITKKY-----IGIRMMSLTSSKAKELKDRHR 403
QY 222 DWGTPPSAIDHCLSVADYDVQVCIHTDTVNEAGYV--DDTLNAMNGRAIHAYHIEGAGG 279
Db 404 DP---PDVI-----SGAYIIEVI--PDTPAEAGGLKENDVIISINGQSV-----VSA 445
QY 280 GHSPDVTMAGEINILPSSTTPTPTTYTINTVAEHL 315
Db 446 NDVSDVIKREBTLNNVVRGNEIDMITV--IPEEID 479
RESULT 14
US-11-013-759-9
; Sequence 9, Application US/11013759
; Publication No. US20050249747A1
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Sasaki, Ken
; APPLICANT: Yang, Yan Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE
; TITLE OF INVENTION: PROTEIN OF MORAXELLA
; FILE REFERENCE: 1038-921MIS-1b
; CURRENT APPLICATION NUMBER: US/11/013,759
; CURRENT FILING DATE: 2004-12-16
; PRIOR APPLICATION NUMBER: US/09/361,619
; PRIOR FILING DATE: 1999-07-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 2053
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-11-013-759-9

Query Match 3.1%; Score 94; DB 7; Length 2053;
Best Local Similarity 19.0%; Pred. No. 13;
Matches 120; Conservative 85; Mismatches 210; Indels 218; Gaps 31;
QY 13 GPTKGDVKRIGDITDLMAEVEHDYTY-----GEELKFGAGKTIR-EGMGQSNPDENT 64
Db 341 GSRDITDAVNAQLKLVLELANRKITFKGDNNSVERGLGNTLTITKGDAOTNALTAN 400
QY 65 LDLV-----ITNAMIIDYTGIVKADI-----GIKNGK- 91
Db 401 IGVTDGNGLVKVLAKELTGLTSVSATNKITVSTNNNAELQSGGLTFSPITGTYDKT 460
QY 92 ---IHG-----IGKAGNKDMQDGVSPHM---VVGVTBALAGE 123
Db 461 VYSIDGLKFTNDSNSIATKGTRITTKKIGFAGTNDGVDESKPYLDNEKLVGNSTLNSG 520
QY 124 GMII--TAGIDSHTHFLSPQFPPTALANGVTTMFGGTGPVDGTNATITPGKWNLHRM 181
Db 521 SLTVNNTTGN-----KQIQVG-ANGI--KEATVANNVANTSATVGT-----558
QY 182 LRAAEYSNMVGFILGKGNSSKKQLV---EOVEAGAIGFKLHEDMGTTPSAIDHCLSVAD 238
Db 559 ARITEE---KIGFAGTNDGVDEQAPYLDKERLKVGRVEI-----TTDSGIN-----AG 603
QY 239 EYDV---QVCIHFD--TVNEAGYVDDTLNAMNGRAIHAYHIEGAGGHSPPDVTMAGEL 292
Db 604 NHKITGLTNGTANTDAVTIKQDKAKPTLNAGDGISINS-----NNGDLVDSGNI 654
QY 293 NILPSSTTPTPTTYTINTV-----AEHLDMLTCHILDKRREDLQFSQ 335
Db 655 -----TTPTVNIISVKTKLNSNGTSGNNKFSVSNADHNNSLVTAADLADYLNKVNETAD 708
QY 336 SRIRPGSIAAEEDVLHDMGVIAWTSSDSQAMGRAGEVIRPTWQTADKNKKEFKLPEDGKD 395
Db 709 SALPSPFKVQNGD-----NSNNAITVGK-----DTNGKTFNTLKLKGN 746
QY 396 NDNPRIKRYISKYTIINPALTHGVSE---YIGSVBECKIADLVVNM--PAFFGVKPKIVIKG 451

Db 747 GVNITTRATGTVTFGIDOSNGLTTPKLVGSDTNGN--RLVIEQVPSADGNSTKNIIG 804
QY 452 -----GMVFSEMGDSNASVPTQPVPYYREMFGHHGKAK-----FDT 488
Db 805 LSPTLPSIASPSGRNIALGNTIIEKDKSNAASIDDLNAGFN--LKNNGKDKDFVSTYDT 862
QY 489 -----SITP-----VSKVAYENGVKKE 505
Db 863 VDFIDGNATTATVTYDEANQTSKVAYDVNVNDEK 895

RESULT 15

US-10-510-386-200
; Sequence 200, Application US/10510386
; Publication No. US20050244922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Olsen, Peter Bjarke
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294.204-US
; CURRENT APPLICATION NUMBER: US/10/510,386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 200
; LENGTH: 1047
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-510-386-200

Query Match 3.1%; Score 93; DB 1; Length 1047;
Best Local Similarity 19.2%; Pred. No. 5.7;
Matches 109; Conservative 75; Mismatches 227; Indels 156; Gaps 27;
QY 42 LKFGAGKTI-----REGMGQNSP--DENTLDLVITN--AMIIDVTGIYKA-DIGIKN 89
Db 562 LKVAASKNVTIVATGNDGVSEISYPASSKYTLISVGATNLDLVSDYSNYGKGLDMVAPG 621
QY 90 GKIHGIGKAGNDQMDOGVs---PHMVVGTEALAGEGMIITAGGIDSHTHFLSPQOFPT 146
Db 622 TDIPSLVPDGNVTYMGSTMAAPHVA-----AAAGLLL-----SQNP SLKPQIAS 667
QY 147 ALANGVTTFMGCGTGPVDGNTATITPGKNLHRLRAAEEYSMNVGFLCKGNSSSKQL 206
Db 668 LLT-----ETTADVAFEQDNPNDYDL 690
QY 207 VEQVEAGAIGFKLHEDWGTTPPSAIDHCLSVADYDVQVCIH----- 247
Db 691 DIEPAAQIPGYDFVSGWGL--NVFHAASV---FELNMKVHPVLNRHTAVTGTA KSGVTV 745
QY 248 -----TDTVNEAGYVDDTLNAMN--GRAIHAYHIEGAGGCHSPD-----VITMAGE 291
Db 746 KILRGQVLGTGTAGKSGAFSVKIPAKAGQVLHV-----AASGHQAETS LRTVVEKAPK 800
QY 292 LNLPSSTTPTIPTYTINTVAEH--LDMLMTCCHLDKRIREDLQFS-----QSRIRPG 341
Db 801 NPSVKRITNKD TAVTGRTAAGYTIKVNACKKVIAQGRADASVSVKVINKQKEYAVLYV 860
QY 342 SIAAEDVLHDMGVIAMTSSDSQAMGRAGEVI PRWTOTADKNKKEFGKLPEDGKDNDFRI 401
Db 861 SASADD--HRESGDVKMTADVIPPFG-----APKVYQVSDKS-----TVIQGKTEANAQV 908
QY 402 KRYISKYITINPALTHGVSEYIGSVEGKIADLVVMNPAFFGVKPKIVIKGMV-----VFS 457
Db 909 SAKAKGTIASGKANGKEYKLSRQKAGTVI-----GVTA K--DKAGNVSKATAVT 959
QY 458 EMGDSNASVPTQPVPYYREMFGHHGKAKFDTSTTFVS--KVAYENGVKKEKG-----LERQ 511
Db 960 VLDKTPPSAPKVPNTVKST-AVKGKAENAAAIIVKSGKKTIGTGKADKKGAPFVKIKKQ 1018

QY 512 VLPVKNCRNITKKDFKEN-DKTAKITV 537
Db 1019 --KANTVLAVTAKDKAGNTSKVSKIY 1043
Search completed: November 28, 2005, 08:25:59
Job time : 9.15365 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 28, 2005, 08:13:38 ; Search time 46.4987 Seconds
(without alignments)
1009.915 Million cell updates/sec

Title: US-09-904-994B-3
Perfect score: 2999
Sequence: 1 MKMKQEVVNTYPTGDKV.....KLCTSKPTSQVPLAQRYTF 568

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues
Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5 COMB.pep.*
2: /cgn2_6/prodata/1/iaa/6 COMB.pep.*
3: /cgn2_6/prodata/1/iaa/H COMB.pep.*
4: /cgn2_6/prodata/1/iaa/PCTUS COMB.pep.*
5: /cgn2_6/prodata/1/iaa/RE COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|----------------------|
| 1 | 2298.5 | 76.6 | 568 | 2 | US-09-338-920B-12 |
| 2 | 2297 | 76.6 | 569 | 2 | US-09-338-920B-6 |
| 3 | 2294 | 76.5 | 569 | 1 | US-08-467-822-26 |
| 4 | 2294 | 76.5 | 569 | 2 | US-08-432-697-26 |
| 5 | 2294 | 76.5 | 569 | 2 | US-08-466-248-26 |
| 6 | 2290 | 76.4 | 566 | 1 | US-08-920-095-3 |
| 7 | 2290 | 76.4 | 566 | 4 | PCT-US96-05800-3 |
| 8 | 2281 | 76.1 | 570 | 2 | US-09-431-705-5 |
| 9 | 2244 | 74.8 | 569 | 1 | US-08-467-822-21 |
| 10 | 2244 | 74.8 | 569 | 2 | US-08-432-697-21 |
| 11 | 2244 | 74.8 | 569 | 2 | US-08-466-248-21 |
| 12 | 2244 | 74.8 | 569 | 2 | US-09-338-920B-10 |
| 13 | 2233 | 74.5 | 559 | 2 | US-09-338-920B-8 |
| 14 | 2028.5 | 67.6 | 568 | 1 | US-07-732-242C-3 |
| 15 | 1828 | 61.0 | 566 | 2 | US-09-252-991A-26887 |
| 16 | 1821 | 60.7 | 625 | 2 | US-09-489-039A-9216 |
| 17 | 1818.5 | 60.6 | 569 | 2 | US-09-543-681A-6029 |
| 18 | 1818 | 60.6 | 567 | 2 | US-09-328-352-5912 |
| 19 | 1817 | 60.6 | 567 | 1 | US-08-967-513-5 |
| 20 | 1817 | 60.6 | 567 | 1 | US-08-687-645B-5 |
| 21 | 1813 | 60.5 | 569 | 1 | US-08-467-822-27 |
| 22 | 1813 | 60.5 | 569 | 2 | US-08-432-697-27 |
| 23 | 1813 | 60.5 | 569 | 2 | US-08-466-248-27 |
| 24 | 1789.5 | 59.7 | 571 | 2 | US-09-710-279-118 |
| 25 | 1788.5 | 59.6 | 573 | 2 | US-09-134-001C-5026 |
| 26 | 1787.5 | 59.6 | 840 | 1 | US-08-467-822-25 |
| 27 | 1787.5 | 59.6 | 840 | 2 | US-08-432-697-25 |

| | | | | | | |
|----|--------|------|-----|---|----------------------|-------------------|
| 28 | 1787.5 | 59.6 | 840 | 2 | US-08-466-248-25 | Sequence 25, Appl |
| 29 | 1708 | 57.0 | 570 | 2 | US-09-602-777A-14 | Sequence 14, Appl |
| 30 | 1619 | 54.0 | 534 | 2 | US-09-602-777A-16 | Sequence 16, Appl |
| 31 | 1166 | 38.9 | 308 | 2 | US-09-338-920B-14 | Sequence 14, Appl |
| 32 | 469 | 15.6 | 121 | 2 | US-10-012-819-160 | Sequence 160, App |
| 33 | 189.5 | 6.3 | 457 | 2 | US-08-415-658-21 | Sequence 21, Appl |
| 34 | 175.5 | 5.9 | 622 | 2 | US-09-949-016-10606 | Sequence 10606, A |
| 35 | 173.5 | 5.8 | 572 | 2 | US-09-949-016-6070 | Sequence 6070, Ap |
| 36 | 173.5 | 5.8 | 578 | 2 | US-09-949-016-11503 | Sequence 11503, A |
| 37 | 171.5 | 5.7 | 572 | 2 | US-09-702-705-1815 | Sequence 1815, Ap |
| 38 | 171.5 | 5.7 | 572 | 2 | US-09-736-457-1815 | Sequence 1815, Ap |
| 39 | 171.5 | 5.7 | 572 | 2 | US-09-671-325-1815 | Sequence 1815, Ap |
| 40 | 171.5 | 5.7 | 572 | 2 | US-10-017-754-1815 | Sequence 1815, Ap |
| 41 | 160.5 | 5.4 | 460 | 1 | US-08-289-709-1 | Sequence 1, Appli |
| 42 | 160.5 | 5.4 | 460 | 1 | US-08-602-656-1 | Sequence 1, Appli |
| 43 | 160 | 5.3 | 507 | 2 | US-09-252-991A-23541 | Sequence 23541, A |
| 44 | 159.5 | 5.3 | 593 | 2 | US-09-949-016-10192 | Sequence 10192, A |
| 45 | 159 | 5.3 | 484 | 2 | US-10-009-782A-2 | Sequence 2, Appli |

ALIGNMENTS

RESULT 1

US-09-338-920B-12
; Sequence 12, Application US/09338920B
; Patent No. 6709851
; GENERAL INFORMATION:
; APPLICANT: Soman, Gopalan
; APPLICANT: Thomas, William D., Jr.
; APPLICANT: Monath, Thomas P.
; TITLE OF INVENTION: Stabilization of Helicobacter Urease
; FILE REFERENCE: 06132/023002
; CURRENT APPLICATION NUMBER: US/09/338,920B
; CURRENT FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: US 08/928,081
; PRIOR FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Helicobacter heilmannii
US-09-338-920B-12

| | | | | | |
|-----------------------|-----|--|---------------------|------------|-------------|
| Query Match | | 76.6% | Score 2298.5; | DB 2; | Length 568; |
| Best Local Similarity | | 73.7% | Pred. No. 7.7e-214; | | |
| Matches 418; | | Conservative | 74; | Mismatches | 74; |
| | | | | Indels | 1; |
| | | | | Gaps | 1; |
| Qy | 2 | KMKKQEVVNTYPTGDKVRLGDTDLWAEVHDYTYGEEELKFGAGKTIREGMGQSNPSD | 61 | | |
| Db | 3 | KISRKEVSMYGPPTGDKVRLGDTDLILEVHDCTTYGEEIKFGGKTIKDMGQTNPS | 62 | | |
| Qy | 62 | ENTLDLVTNAMIIDYTGIKADIGIKNGKIHGKAGNKMDQGVSPHVVGVGTEALA | 121 | | |
| Db | 63 | SHELDLVTNALIVDYGIVKADIGIKNGKIHGKAGNKLDQGVNRLCVGPATEALA | 122 | | |
| Qy | 122 | EGEMITAGDISHTFLSPQOFTALANGVTTFGGTGPVDTGNATITPKKNLHRM | 181 | | |
| Db | 123 | AEGLIVTAGGIDTHIFISPOQIPTAFASGITTTMGGTGPDAGTGNATITPGRNLRKM | 182 | | |
| Qy | 182 | LRAEEYSMVNVLGKGNSSKKQLVQVEAGAIGFKLHEDWGTTPSAIDHCLSVADYD | 241 | | |
| Db | 183 | LRASEYAMNLGYLGKGNVSPFALIDOLEAGAIGFKLHEDWGTTPSAINHALNTADKYD | 242 | | |
| Qy | 242 | VQVCIHDTVNEAGYVDDTLNANNGRAIHAYHIEGAGGCHSPDVTITMAGELNILPSSTTP | 301 | | |
| Db | 243 | VQVAIHTDLNEACGVEDTLEAIAGRTHHTHTEGAGGCHAPDVIKMAEFNILLASTNP | 302 | | |
| Qy | 302 | TIPYTINTVAEHLDMLTCHHLDKRREDLOFSRIRPGSIAEDVLHDMGVIAWTSSD | 361 | | |
| Db | 303 | TIPFTKTEAEHMDMLM-CHHLDNKXIEDVEFADSRIRPQTIAARDKLHDMGIFITSDD | 361 | | |

QY 362 SOAMGRAGEVIRTRWQTADKNKBERGKLPEDGKONDNDNFRIKRYISKYITINPALTHGVSEY 421
DB 362 SOAMGRVGEVIRTRWQTADKNKBERGLPEEXGDNDFRIKRYISKYITINPAITHGISEY 421
QY 422 IGSVEEGKADLAVNPAFFGVKPKIVIKGGMVSEMGDSNASVPTPOPVYREMGHH 481
DB 422 VGSVEVGKADLVNWSPAFFGKIPNMIIKGGFIALSQMGDANASIPTPOPVYREMGHH 481
QY 482 GKAKEDTSTITFVSKVAYENGVEKELGLERQVLPVKNCRNITKKDFKFNDDTKAKITVDPKT 541
DB 482 GKAKEDTNTITFVSQVAYENGIRHELGLQVLPVKNCRNITKKDLKFNDVTAHIEVNPET 541
QY 542 FEVFDGKCLTSKPTSQVPLAORYTFF 568
DB 542 YKVKVDGNEVTSAAADKLSLAQLYNLF 568
RESULT 2
US-09-338-920B-6
; Sequence 6, Application US/09338920B
; Patent No. 6709851
; GENERAL INFORMATION:
; APPLICANT: Soman, Gopalan
; APPLICANT: Thomas, William D., Jr.
; TITLE OF INVENTION: Stabilization of Helicobacter Urease
; FILE REFERENCE: 06132/023002
; CURRENT APPLICATION NUMBER: US/09/338,920B
; CURRENT FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: US 08/928,081
; PRIOR FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-338-920B-6

Query Match 76.6%; Score 2297; DB 2; Length 569;
Best Local Similarity 73.4%; Pred. No. 1.le-213;
Matches 416; Conservative 76; Mismatches 75; Indels 0; Gaps 0;
QY 2 KMKQBYNTYGTGDKVRLGDTDLWAEVHDYTYGELKFGAGKTIREGMQSNPSD 61
DB 3 KISRKEYVSMYGTGDKVRLGDTDLIAVEHDYTYGELKFGAGKTIREGMQSNPS 62
QY 62 ENTLDLVITNAMIIDYTYGIKADIGIKNGKIHGKAGNKMDDGVSHPMVVGTEALA 121
DB 63 KEELDLIITNALIVDVTGIYKADIGIKDKGIAGKGNKMDQGVKNLSVGPATEALA 122
QY 122 GEGMIITAGGIDSHTHFLSPQFPPTALANGVTMTFGGGTGPVDGNTATTITPGKWNLM 181
DB 123 GEGLIVTAGGIDTHIFISQOIPPTAFASGVTTMIGGGTGPADGNTATTITPGRRLKWM 182
QY 182 LRAAEYSNMVGLFGKNGSSKQVQEAIGKPLHEDGTTPSAIDCHLSVADEYD 241
DB 183 LRAAEYSNMVGLFLAGNASNDASLADQIEAGAIGKPLHEDGTTPSAINHLDVADKYD 242
QY 242 VQVCIHDTVNEAGYVDDTLNANNGRAIHAYHIEGAGGHSPTVITMAGELNLPSTTP 301
DB 243 VQVAIHDTLNEAGCVEDTMAAIGRTHMTFTHEGAGGHPDIIKVAGEHNLIPASTNP 302
QY 302 TIPTYINTVAHLDMLMTCCHLDKXIREDLQFSQSRIRPGSIAAEVLDHDMGVIAWSSD 361
DB 303 TTPFTVNTAEHMDMLVMVCHLDKSKEDVQFADSRIRPTIAAEDTLHDMGIFSTSSD 362
QY 362 SOAMGRAGEVIRTRWQTADKNKBERGKLPEDGKONDNDNFRIKRYISKYITINPALTHGVSEY 421
DB 363 SOAMGRVGEVIRTRWQTADKNKBERGLPEEXGDNDFRIKRYISKYITINPAITHGISEY 422
QY 422 IGSVEEGKADLAVNPAFFGVKPKIVIKGGMVSEMGDSNASVPTPOPVYREMGHH 481

DB 423 VGSVEVGKADLVNWSPAFFGVKPNMIIKGGFIALSQMGDANASIPTPOPVYREMGHH 482
QY 482 GKAKEDTSTITFVSKVAYENGVEKELGLERQVLPVKNCRNITKKDFKFNDDTKAKITVDPKT 541
DB 483 GKAKYDANITFVSQAAYDKIKELGLERQVLPVKNCRNITKKDMQFNDDTAHIEVNPET 542
QY 542 FEVFDGKCLTSKPTSQVPLAORYTFF 568
DB 543 YHVFVDGKEVTSKANKVSLAQLFSIF 569
RESULT 3
US-08-467-822-26
; Sequence 26, Application US/08467822
; Patent No. 5843460
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,822
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/447,177
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 569 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-467-822-26

Query Match 76.5%; Score 2294; DB 1; Length 569;
Best Local Similarity 73.4%; Pred. No. 2.le-213;
Matches 416; Conservative 76; Mismatches 75; Indels 0; Gaps 0;
QY 2 KMKQBYNTYGTGDKVRLGDTDLWAEVHDYTYGELKFGAGKTIREGMQSNPSD 61
DB 3 KISRKEYVSMYGTGDKVRLGDTDLIAVEHDYTYGELKFGAGKTIREGMQSNPS 62


```
;
; ADDRESS: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,248
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/447,177
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 569 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-466-248-26

Query Match 76.5%; Score 2294; DB 2; Length 569;
Best Local Similarity 73.4%; Pred. No. 2.1e-213; Indels 0; Gaps 0;
Matches 416; Conservative 76; Mismatches 75;

QY 2 KMKQEVNTYGTGDKVRLGDTDLWAEVHDYTYGEEELKFGAGKTIRGGMGQSNPD 61
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 3 KISRKEYVSMYGTGDKVRLGDTDLIAEVHDYTYGEEELKFGAGKTIRGGMGQSNPS 62
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 62 ENTLDLVITNAMIIDYTYGIYKADIGIKNGKIHGIGKAGNKMDQGVSPHMVVGVPTEALA 121
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 63 KEELDLIITNALIVDYTYGIYKADIGIKNGKIHGIGKAGNKMDQGVSPHMVVGVPTEALA 122
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 122 GEGMIITAGGIDSHTHFLSPQFPPTALANGVTMTMFGGGTGPVDGTNATTITPGKNLHRM 181
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 123 GEGLIIVTAGGIDTHHFIHQIPQIPAFASGVTTMIGGGTGPADGTNATTITPGRRNLKWM 182
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 182 LRAAEYSNMVGLFGKNSKKQLVEQVAGAIKFKLHEDWGTTPSAIDHCLSVADYD 241
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 183 LRAAEYSNMVGLFLAGNASNASARDQIEAGAIKFKIHEDWGTTPSAINHALDVADKYD 242
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 242 VQVCIHTDVTNAGYVDDTLNAMNGRAIHAYHIEGAGGGHSPDVTMAGELNLPSTTP 301
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 243 VQVAIHTDVTNAGCVEDTMAIAGRTMHTFTEGAGGGHAPDIKVAEGHNILPASTNP 302
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 302 TTPYINTVAEHLDMTMCHLDKRLREDLQFSQSRIRGSIAAEDVLHDMGVAMTSSD 361
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 303 TTPFTVNTAEHMDMLVMVCHLDKSLKEDVQFADSRIRPQTAAEDTLHDMGIFSTSSD 362
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 362 SQAMGRAGVIRPTWQTADKNKKEFGKLPEDGKNDNFRKRYISKYITNPALTCHGVSEY 421
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 363 SQAMGRVGVIRPTWQTADKNKKEFGKLPEDGKNDNFRKRYISKYITNPALTCHGVSEY 422
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 422 IGSVEEGKIADLVWNPAPFGVGPVKPVIKGMVVFSEMGDSNASVPTPQPVYVREMFQGH 481
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

;
; ADDRESS: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,248
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/447,177
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 569 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-466-248-26

Query Match 76.4%; Score 2290; DB 1; Length 566;
Best Local Similarity 73.6%; Pred. No. 5.1e-213;
Matches 415; Conservative 75; Mismatches 74; Indels 0; Gaps 0;

QY 2 KMKQEVNTYGTGDKVRLGDTDLWAEVHDYTYGEEELKFGAGKTIRGGMGQSNPD 61
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 3 KISRKEYVSMYGTGDKVRLGDTDLIAEVHDYTYGEEELKFGAGKTIRGGMGQSNPS 62
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 62 ENTLDLVITNAMIIDYTYGIYKADIGIKNGKIHGIGKAGNKMDQGVSPHMVVGVPTEALA 121
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 63 KEELDLIITNALIVDYTYGIYKADIGIKNGKIHGIGKAGNKMDQGVSPHMVVGVPTEALA 122
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 122 GEGMIITAGGIDSHTHFLSPQFPPTALANGVTMTMFGGGTGPVDGTNATTITPGKNLHRM 181
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 123 GEGLIIVTAGGIDTHHFIHQIPQIPAFASGVTTMIGGGTGPADGTNATTITPGRRNLKWM 182
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 182 LRAAEYSNMVGLFGKNSKKQLVEQVAGAIKFKLHEDWGTTPSAIDHCLSVADYD 241
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 183 LRAAEYSNMVGLFLAGNASNASARDQIEAGAIKFKIHEDWGTTPSAINHALDVADKYD 242
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

;
; ADDRESS: Cynthia K. Lee et al.
; TITLE OF INVENTION: MULTIMERIC, RECOMBINANT UREASE VACCINE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,095
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/431,041
; FILING DATE: 28-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06132/013001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 566 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-920-095-3
```

```
QY 242 VQVCIHTDTVNEAGYVDDTLNMMNGRAIHAYHIEGAGGGHSPDVITWAGELNLPSTTP 301
Db 243 VQVAIHDTLNEAGCVEDTMAAIAGRTHMTHTFTGEGAGGHPDIIKVAGEHNILPASTNP 302
QY 302 TTPYTTINTVAEHLDMLTCHHLDKRIREDLQFSQSRIKPGSIAAEDVLHDMGVIAWTSDD 361
Db 303 TTPFTVNTAEHMDMLVCHHLDSKIKEDVQFADSRIRPOTIAAEDTLHDMGIFSTSSD 362
QY 362 SQAMGRAGEVIRPTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYTTINPALTHGVSEY 421
Db 363 SQAMGRAGEVIRPTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYTTINPALTHGVSEY 422
QY 422 IGSVEGKIADLVVNNPAFFGVKPKIVIKGMVVFSEMGDSNASVPTPOPVYVREMGHH 481
Db 423 VGSVEGKVADLVWSPAFFGVKPNMIIKGGFIALSQMGDANASIPTPOPVYVREMFHH 482
QY 482 GKAKPDTSTFVSQVAYENGVEKGLERQVLVKNCRNITKDKFENDKTAKITVDPKT 541
Db 483 GKAKYDANITFVSQAAYDKGKEBELGLERQVLVKNCRNITKDKMQFNDDTAHIEVNPET 542
QY 542 FEVFDGKLTCKTSKPTSQVPLAQR 565
Db 543 YHVFVDGKEVTSKPANKVSLAQLF 566

RESULT 7
PCT-US96-05800-3
; Sequence 3, Application PC/TUS9605800
; GENERAL INFORMATION:
; APPLICANT: Oravax, Inc.
; TITLE OF INVENTION: MULTIMERIC, RECOMBINANT UREASE VACCINE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/05800
; FILING DATE: 23-APR-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/431,041
; FILING DATE: 28-APR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/568,122
; FILING DATE: 06-DEC-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06132/020001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 566 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US96-05800-3

Query Match 76.4%; Score 2290; DB 4; Length 566;
Best Local Similarity 73.6%; Pred. No. 5.1e-213;
Matches 414; Conservative 76; Mismatches 77; Indels 0; Gaps 0;
```

```
Matches 415; Conservative 75; Mismatches 74; Indels 0; Gaps 0;
QY 2 KMKKQEVYNTYGPDKGKVRIGDITDLWAEVHDYTYTYGEEFKFGAGKTIREGNGSQNSPD 61
Db 3 KISRKEYVSMYGPPTGDKVRLGDTDLIAEVEHDYTYTYGEEFKFGAGKTIREGNGSQNSPD 62
QY 62 ENTLDLVITNAMIIDYTYGIYKADIGIKNGKIHGIGKAGNKMDDGVSPHVVGVGTTEALA 121
Db 63 KEELDILITNALIVDYTYGIYKADIGIKNGKIHGIGKAGNKMDDGVSPHVVGVGTTEALA 122
QY 122 GEGMIITAGGIDSTHFLSPQFPPTALANGVTTFMGGGTGPVDGTNATTTTGGKMLHRM 181
Db 123 GEGMIITAGGIDSTHFLSPQFPPTALANGVTTFMGGGTGPVDGTNATTTTGGKMLHRM 182
QY 182 LRAAEYSMNVGFLGKGNSSKKOLVEQVEAGAIGFKLHEDMGTTPTSAIDHCLSLVADEYD 241
Db 183 LRAAEYSMNLGFLAKGNASNDASLADQIEAGAIGFKLHEDMGTTPTSAIDHCLSLVADEYD 242
QY 242 VQVCIHTDTVNEAGYVDDTLNMMNGRAIHAYHIEGAGGGHSPDVITWAGELNLPSTTP 301
Db 243 VQVAIHDTLNEAGCVEDTMAAIAGRTHMTHTFTGEGAGGHPDIIKVAGEHNILPASTNP 302
QY 302 TTPYTTINTVAEHLDMLTCHHLDKRIREDLQFSQSRIKPGSIAAEDVLHDMGVIAWTSDD 361
Db 303 TTPFTVNTAEHMDMLVCHHLDSKIKEDVQFADSRIRPOTIAAEDTLHDMGIFSTSSD 362
QY 362 SQAMGRAGEVIRPTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYTTINPALTHGVSEY 421
Db 363 SQAMGRAGEVIRPTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYTTINPALTHGVSEY 422
QY 422 IGSVEGKIADLVVNNPAFFGVKPKIVIKGMVVFSEMGDSNASVPTPOPVYVREMGHH 481
Db 423 VGSVEGKVADLVWSPAFFGVKPNMIIKGGFIALSQMGDANASIPTPOPVYVREMFHH 482
QY 482 GKAKPDTSTFVSQVAYENGVEKGLERQVLVKNCRNITKDKFENDKTAKITVDPKT 541
Db 483 GKAKYDANITFVSQAAYDKGKEBELGLERQVLVKNCRNITKDKMQFNDDTAHIEVNPET 542
QY 542 FEVFDGKLTCKTSKPTSQVPLAQR 565
Db 543 YHVFVDGKEVTSKPANKVSLAQLF 566

RESULT 8
US-09-431-705-5
; Sequence 5, Application US/09431705
; Patent No. 6585975
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Londono-Arcila, Patricia
; APPLICANT: Freeman, Donna
; TITLE OF INVENTION: Use of salmonella vectors for
; TITLE OF INVENTION: vaccination against helicobacter infection
; FILE REFERENCE: 06132/060001
; CURRENT APPLICATION NUMBER: US/09/431,705
; CURRENT FILING DATE: 1999-11-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 570
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-431-705-5

Query Match 76.1%; Score 2281; DB 2; Length 570;
Best Local Similarity 73.0%; Pred. No. 3.9e-212;
Matches 414; Conservative 76; Mismatches 77; Indels 0; Gaps 0;
QY 2 KMKKQEVYNTYGPDKGKVRIGDITDLWAEVHDYTYTYGEEFKFGAGKTIREGNGSQNSPD 61
Db 4 KISRKEYVSMYGPPTGDKVRLGDTDLIAEVEHDYTYTYGEEFKFGAGKTIREGNGSQNSPD 63
QY 62 ENTLDLVITNAMIIDYTYGIYKADIGIKNGKIHGIGKAGNKMDDGVSPHVVGVGTTEALA 121
```

Db 64 KEELELITNALIVDTYGIYKADIGIKGKIAGIKGKNKMDQGVKNLNSVGPATEALA 123
QY 122 GEGMIITAGGIDSHTHFLSPQOFTALANGVTTMFGGTGPVDGTNATITPGKNLHRM 181
Db 124 GEGLIIVTAGGIDTHIHFIISPQIPTAFASGVTTMIGGTGPADGTNATITPGRENKWM 183
QY 182 LRAAEYSMNVLGKGNSSSKQLVQVEAGAIGFKLHEDMGTPSPAIHCLSVADRYD 241
Db 184 LRAAEYSMNVLGKGNSSSKQLVQVEAGAIGFKLHEDMGTPSPAIHCLSVADRYD 243
QY 242 VOVCIHDTVNEAGVDDTLNAMNGRAIHAYHIEGAGGSHSDVITMAGELNIPSSITP 301
Db 244 VOVAIATDTLNSAGVEDTMAAIGRTHTFTTEGAGGSHSDVITMAGELNIPSSITP 303
QY 302 TIPTTINTVAEHLDMTCHLHDKREIDLPQSQRIRPGSIARSDVILHDMGVIAHTSSD 361
Db 304 TIPTTINTVAEHLDMTCHLHDKREIDLPQSQRIRPGSIARSDVILHDMGVIAHTSSD 363
QY 362 SOAMGRAGEVIRPTWOTADKNKKEFGKLPEDGKNDNFRIKRYISKYTTINPALTGVSEY 421
Db 364 SOAMGRAGEVIRPTWOTADKNKKEFGKLPEDGKNDNFRIKRYISKYTTINPALTGVSEY 423
QY 422 IGSVEEGKIADLVNPNPAFFGVKPKIVIKGGMVFSMGDSNASVPTPQPVYVREMFGHH 481
Db 424 VGSVEVGKADLVNPNPAFFGVKPKIVIKGGMVFSMGDSNASVPTPQPVYVREMFGHH 483
QY 482 GKAKFDTSITFVSKVAYENGVEKGLGLERQVLVKNCRNITKKDKFNDKTAKITVDPKT 541
Db 484 GKAKFDTSITFVSKVAYENGVEKGLGLERQVLVKNCRNITKKDKFNDKTAKITVDPKT 543
QY 542 FEVFDGKLCSTSKPTSOVPLAQRYTFF 568
Db 544 YHVFVDGKEVTSKPNKVSQAQYLSIF 570

RESULT 9

US-08-467-822-21
; Sequence 21, Application US/08467822
; Patent No. 5843460
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESSES:
; ADDRESSES: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSES: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,822
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/447,177
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/432,697

; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 569 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..569
; OTHER INFORMATION: /note= "URE B - FIGURE 3."
US-08-467-822-21
Query Match 74.8%; Score 2244; DB 1; Length 569;
Best Local Similarity 72.5%; Pred. No. 1.5e-208; Mismatches 85; Indels 0; Gaps 0;
Matches 411; Conservative 71;
QY 2 KKKKQBYVNTYGTGDKVRLGDTDLMAEVEHDYTTYGEELKFGAGKTIREGMQSNSPD 61
Db 3 KISKEVSMVSGPTTGDRVRLGDTDLILEVHEDCTTYGEEIKFGGKTIRDMGQSNPS 62
QY 62 ENTLDLVITNMIIDYTYGIYKADIGIKGKAGNKMDDQGVSPHMVVGTEALA 121
Db 63 SYELDLVITNALIVDTYGIYKADIGIKGKAGNKMDDQGVSPHMVVGTEALA 122
QY 122 GEGMIITAGGIDSHTHFLSPQOFTALANGVTTMFGGTGPVDGTNATITPGKNLHRM 181
Db 123 AGLIVTAGGIDTHIHFIISPQIPTAFASGVTTMIGGTGPADGTNATITPGRENKWM 182
QY 182 LRAAEYSMNVLGKGNSSSKQLVQVEAGAIGFKLHEDMGTPSPAIHCLSVADRYD 241
Db 183 LRAAEYSMNVLGKGNSSSKQLVQVEAGAIGFKLHEDMGTPSPAIHCLSVADRYD 242
QY 242 VOVCIHDTVNEAGVDDTLNAMNGRAIHAYHIEGAGGSHSDVITMAGELNIPSSITP 301
Db 243 VOVAIATDTLNSAGVEDTMAAIGRTHTFTTEGAGGSHSDVITMAGELNIPSSITP 302
QY 302 TIPTTINTVAEHLDMTCHLHDKREIDLPQSQRIRPGSIARSDVILHDMGVIAHTSSD 361
Db 303 TIPTTINTVAEHLDMTCHLHDKREIDLPQSQRIRPGSIARSDVILHDMGVIAHTSSD 362
QY 362 SOAMGRAGEVIRPTWOTADKNKKEFGKLPEDGKNDNFRIKRYISKYTTINPALTGVSEY 421
Db 363 SOAMGRAGEVIRPTWOTADKNKKEFGKLPEDGKNDNFRIKRYISKYTTINPALTGVSEY 422
QY 422 IGSVEEGKIADLVNPNPAFFGVKPKIVIKGGMVFSMGDSNASVPTPQPVYVREMFGHH 481
Db 423 VGSVEVGKADLVNPNPAFFGVKPKIVIKGGMVFSMGDSNASVPTPQPVYVREMFGHH 482
QY 482 GKAKFDTSITFVSKVAYENGVEKGLGLERQVLVKNCRNITKKDKFNDKTAKITVDPKT 541
Db 483 GKAKFDTSITFVSKVAYENGVEKGLGLERQVLVKNCRNITKKDKFNDKTAKITVDPKT 542
QY 542 FEVFDGKLCSTSKPTSOVPLAQRYTFF 568
Db 543 YHVFVDGKEVTSKPNKVSQAQYLSIF 569
RESULT 10
US-08-432-697-21
; Sequence 21, Application US/08432697
; Patent No. 6248330
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes

```

; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 569 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..569
; OTHER INFORMATION: /note= "URE B - FIGURE 3."
; US-08-432-697-21

Query Match 74.8%; Score 2244; DB 2; Length 569;
Best Local Similarity 72.5%; Pred. No. 1.5e-208;
Matches 411; Conservative 71; Mismatches 85; Indels 0; Gaps 0;

Qy 2 KMKQBYVNTYPTKDKVRLGDTDLAWEVHDYTYTGEEELKFGAGKTIREGMQSNSPD 61
Db 3 KISRKEVSVMYGPTTGDVRLGDTDLILEVHDCTTYGEEIKFGGKTIIRDGMSQTNSPS 62
Qy 62 ENTLDLVITNMIIDYTGIIYKADIGIKNGKIHGIKAGNKQMDGVSPHMVGVGTEALA 121
Db 63 SYELDLVLITNALIVDYGIIYKADIGIKDKGIAGIKAGNKQMDQGVNNDLCVGPATEALA 122
Qy 122 GEGMIIITAGIDISHTHFLSPQPPFALANGVTMTFGGGTGPVDGTNATTITPGKNLHRM 181
Db 123 AEGLIVTAGIDITHIHFISFPQIIPAFASGVTMTLGGGTGPADGTNATTITPGRANLKSM 182
Qy 182 LRAAEYSNMVGLFGKNGSSKKQLVEQVEAGAIGFKLHEDGTTTPSAIDHCLSVADEYD 241
Db 183 LRAAEHYANMLGFLAKGNVSYEPSLEDQIEAGAIQFKIHEDWGSTPAIHHCLNVADEYD 242
Qy 242 VQVCIIHTDVNEAGYVDDTLNAMNGRAIHAYHIEGAGGCHSPDVTIMAGELNLPSTTP 301
Db 243 VQVAIHDTLNEAGCVEDTLEIAGRTIHTFHTEGAGGCHAPDVIKWAGEFNILPASTNP 302
Qy 302 TIPTYTNTVAEHLDMLTCHLHDKRIREDLQPSQSRIRPGSIAAEDVLHDMGVIAMTSSD 361
Db 303 TIFPTKNTAEHMDMLMVCHLHDKSIKEVDQFADSRIRPQTAAEDQLHDMGIFITSSD 362

```

OTHER INFORMATION: /note= "URE B - FIGURE 3."

US-08-466-248-21

Query Match 74.8%; Score 2244; DB 2; Length 569;
 Best Local Similarity 72.5%; Pred. No. 1.5e-208;
 Matches 411; Conservative 71; Mismatches 85; Indels 0; Gaps 0;

QY 2 KMKKQEVNTYGTGDKVRLGDTDLWAEVEHDYTTYGBELKFGAGKTTIREGMSQNSPD 61
 DB 3 KISRKEYVSMYGTGDRVRLGDTDLILEVEHDCTTYGBEIKFEGGKTTIRDCGMSQNSPS 62

QY 62 ENTLDLIVTNAMIIDYTGKADIGIKNGKIHGIGKAGNKMOMDGVSPHVMVGVGTEALA 121
 DB 63 SYELDLVLNALLIVDYGKADIGIKGKIAGIGKAGNKMOMDGVNNLVCVPATEALA 122

QY 122 GEGMIITAGGIDSHTHFLSPQPPFTALANGVTTMTFGGTCGPVDGTNATTTTPGKNLHRM 181
 DB 123 AEGLVITAGGIDTHIHFIISPOQIPTAFASGVTTMIGGTCGPADGTNATTTTPGRANLKSM 182

QY 182 LRAAEYSNMVGLFGKGNSSKKQLVEQVAGAGFKLHEDWGTTPSAIDHCLUSVADEYD 241
 DB 183 LRAAEYAMNLGFLAKGNVSEPSLRDQIEAGAGFKIHEDWGSTPAAIHHCLNVADEYD 242

QY 242 VOVCITHTDVNEAGYVDDTLNAMNGRAIHAYHIEGAGGSHSPDVI TWAGELNLPSTTP 301
 DB 243 VQVAIHTDTLINEAGCVEDTLEAIGRTIHTFHTGAGGGHAPDVI KWAGEFNILPASTNP 302

QY 302 TIPTYTINTVAEHLDMMLTCHHLDKRIRDLQFSQSIRPGSIAAEDVLHDMGVIAMTSSD 361
 DB 303 TIPTKNTAEHMDMLVCHLHDKSIEDVQADSRIRPQTIAAEQLHDMGIFITSDD 362

QY 362 SQAMGRAGEVIRPTWOTADKNKKEFGKLPEDGKNDNFRIKRYISKYTTINPALTHGVSEY 421
 DB 363 SQAMGRVGEVIRPTWOTADKNKKEFGRLKEEGKNDNFRIKRYISKYTTINPGIAHGISDY 422

QY 422 IGSVEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPOPVYVREMGHH 481
 DB 423 VGSVEVGKYADLVWSPAFPGIKPNMIIKGGFTALSQMGDANASIPTPOPVYVREMGHH 482

QY 482 GKAKFDTSTITFVSKVAYENGVEKGLGLERQVLPVKNCRNITTKDKFNDKTAKITVDPKT 541
 DB 483 GKKNKFTNTITFVSOAYKAGIKBELGLDRAAPPVKNCNITTKDLKFNVDVAHIDVNPET 542

QY 542 FEYVVDGKLCSTKPTSQVPLAQRYTFF 568
 DB 543 YKVKVDGKEVTSKAADELSLAQLYNLF 569

RESULT 13

US-09-338-920B-8
 ; Sequence 8, Application US/09338920B
 ; Patent No. 6709851
 ; GENERAL INFORMATION:
 ; APPLICANT: Soman, Gopalan
 ; APPLICANT: Thomas, William D., Jr.
 ; APPLICANT: Monath, Thomas P.
 ; TITLE OF INVENTION: Stabilization of Helicobacter Urease
 ; FILE REFERENCE: 06132/023002
 ; CURRENT APPLICATION NUMBER: US/09/338,920B
 ; CURRENT FILING DATE: 2002-12-23
 ; PRIOR APPLICATION NUMBER: US 08/928,081
 ; PRIOR FILING DATE: 1997-09-12
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: Fast-SEQ for Windows Version 4.0
 ; SEQ ID NO 8
 ; LENGTH: 559
 ; TYPE: PRT
 ; ORGANISM: Helicobacter pylori
 ; US-09-338-920B-8

Query Match 74.5%; Score 2233; DB 2; Length 559;
 Best Local Similarity 73.5%; Pred. No. 1.7e-207;
 Matches 405; Conservative 70; Mismatches 76; Indels 0; Gaps 0;

QY 2 KMKKQEVNTYGTGDKVRLGDTDLWAEVEHDYTTYGBELKFGAGKTTIREGMSQNSPD 61
 DB 3 KISRKEYVSMYGTGDRVRLGDTDLIAEVEHDYTTYGBELKFGGKTTIREGMSQNSPS 62

QY 62 ENTLDLIVTNAMIIDYTGKADIGIKNGKIHGIGKAGNKMOMDGVSPHVMVGVGTEALA 121
 DB 63 SYELDLVLNALLIVDYGKADIGIKGKIAGIGKAGNKMOMDGVNNLVCVPATEALA 122

QY 122 GEGMIITAGGIDSHTHFLSPQPPFTALANGVTTMTFGGTCGPVDGTNATTTTPGKNLHRM 181
 DB 123 AEGLVITAGGIDTHIHFIISPOQIPTAFASGVTTMIGGTCGPADGTNATTTTPGRANLKSM 182

QY 182 LRAAEYSNMVGLFGKGNSSKKQLVEQVAGAGFKLHEDWGTTPSAIDHCLUSVADEYD 241
 DB 183 LRAAEYAMNLGFLAKGNVSEPSLRDQIEAGAGFKIHEDWGSTPAAIHHCLNVADEYD 242

QY 242 VOVCITHTDVNEAGYVDDTLNAMNGRAIHAYHIEGAGGSHSPDVI TWAGELNLPSTTP 301
 DB 243 VQVAIHTDTLINEAGCVEDTLEAIGRTIHTFHTGAGGGHAPDVI KWAGEFNILPASTNP 302

QY 302 TIPTYTINTVAEHLDMMLTCHHLDKRIRDLQFSQSIRPGSIAAEDVLHDMGVIAMTSSD 361
 DB 303 TIPTKNTAEHMDMLVCHLHDKSIEDVQADSRIRPQTIAAEQLHDMGIFITSDD 362

QY 362 SQAMGRAGEVIRPTWOTADKNKKEFGKLPEDGKNDNFRIKRYISKYTTINPALTHGVSEY 421
 DB 363 SQAMGRVGEVIRPTWOTADKNKKEFGRLKEEGKNDNFRIKRYISKYTTINPGIAHGISDY 422

QY 422 IGSVEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPOPVYVREMGHH 481
 DB 423 VGSVEVGKYADLVWSPAFPGIKPNMIIKGGFTALSQMGDANASIPTPOPVYVREMGHH 482

QY 482 GKAKFDTSTITFVSKVAYENGVEKGLGLERQVLPVKNCRNITTKDKFNDKTAKITVDPKT 541
 DB 483 GKKNKFTNTITFVSOAYKAGIKBELGLDRAAPPVKNCNITTKDLKFNVDVAHIDVNPET 542

QY 542 FEYVVDGKLCSTKPTSQVPLAQRYTFF 568
 DB 543 YKVKVDGKEVTSKAADELSLAQLYNLF 569

RESULT 12

US-09-338-920B-10
 ; Sequence 10, Application US/09338920B
 ; Patent No. 6709851
 ; GENERAL INFORMATION:
 ; APPLICANT: Soman, Gopalan
 ; APPLICANT: Thomas, William D., Jr.
 ; APPLICANT: Monath, Thomas P.
 ; TITLE OF INVENTION: Stabilization of Helicobacter Urease
 ; FILE REFERENCE: 06132/023002
 ; CURRENT APPLICATION NUMBER: US/09/338,920B
 ; CURRENT FILING DATE: 2002-12-23
 ; PRIOR APPLICATION NUMBER: US 08/928,081
 ; PRIOR FILING DATE: 1997-09-12
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: Fast-SEQ for Windows Version 4.0
 ; SEQ ID NO 10
 ; LENGTH: 569
 ; TYPE: PRT
 ; ORGANISM: Helicobacter felis
 ; US-09-338-920B-10

Query Match 74.8%; Score 2244; DB 2; Length 569;
 Best Local Similarity 72.5%; Pred. No. 1.5e-208;
 Matches 411; Conservative 71; Mismatches 85; Indels 0; Gaps 0;

Db 63 KEELDLITNALIVDTYGIYKADIGIKGAGIKGGNKDQDGKYNLNVGVPATEALA 122
Qy 122 GEGMIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGGTGPVDGNTNATTITPGKWNLRHM 181
Db 123 GEGLIIVTAGGIDTHHFTISPOIPTAFASGVTTMGGGTGPADGTNATTITPGRNLRKM 182
Qy 182 LRAAEYSNMVFLGKGNSSKKQLVQVEAGAIGFKLHEDWGTTPSAIDHCLSVADRYD 241
Db 183 LRAAEYSNMVFLGKGNSSKKQLVQVEAGAIGFKLHEDWGTTPSAIDHCLSVADRYD 242
Qy 242 VQVCHTDTVNEAGYVDDTLNANWGRAIHAYHIEGAGGSHSPDVTMAGELNLPSSSTP 301
Db 243 VQVAIHDTTDLNAGCVEDTMAAIAGRTHMTYHTEGAGGSHAPDIIVKAGEHNILPASTNP 302
Qy 302 TIPTYTINTVAEHLDMMLTCHLDKRIREDLOFSQSRIKRGSIKRYISKYTINPALTHGVSEY 361
Db 303 TIPTYTINTVAEHLDMMLTCHLDKRIREDLOFSQSRIKRGSIKRYISKYTINPALTHGVSEY 362
Qy 362 SOAMGRAGEVIRPTWQTADKNKKEFKLPEDGKNDNFRIKRYISKYTINPALTHGVSEY 421
Db 363 SOAMGRAGEVIRPTWQTADKNKKEFKLPEDGKNDNFRIKRYISKYTINPALTHGVSEY 422
Qy 422 IGSVEEGKIADLVVWNPAPFGVKPILVKGGMVVFSEMGDSNASVPTPQPVYRPMFGHH 481
Db 423 VGSVEVGKVLADLVWNPAPFGVKPILVKGGMVVFSEMGDSNASVPTPQPVYRPMFAHH 482
Qy 482 GKAKEDTSTITFVSKVAYENGVEKGLERQVLVKNCRNITKKDFKFNDKTAKITVDPKT 541
Db 483 GKAKEDTSTITFVSKVAYENGVEKGLERQVLVKNCRNITKKDFKFNDKTAKITVDPKT 542
Qy 542 FEVFDGKLCIT 552
Db 543 YHVFVDGKEVT 553

RESULT 14

US-07-732-242C-3
; Sequence 3, Application US/07732242C
; Patent No. 5298399
; GENERAL INFORMATION:
; APPLICANT: Uozumi, Takeashi; Masaki, Haruhiko;
; APPLICANT: Hidaka, Makoto; Nakamura, Akira;
; APPLICANT: Maeda, Michihisa; Yoneta, Yasuo
; TITLE OF INVENTION: Gene of Urease
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Frishauf, Holtz, Goodman & Woodward, P.C.
; STREET: 600 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10016-2088
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 0.72mb
; COMPUTER: IBM PC compatible (NEC PC-9801 RX)
; OPERATING SYSTEM: MS DOS
; SOFTWARE: ASCII Form
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/732,242C
; FILING DATE: 19910718
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JPN 2-210178
; FILING DATE: 10-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Goodman, Herbert
; REGISTRATION NUMBER: 17081
; REFERENCE/DOCKET NUMBER: 910532/HG
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)972-1400
; TELEFAX: (212)370-1622
; TELEX: 236268

; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 568 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-07-732-242C-3

Query Match 67.6%; Score 2028.5; DB 1; Length 568;
Best Local Similarity 66.1%; Pred. No. 1.2e-187;
Matches 375; Conservative 77; Mismatches 112; Indels 3; Gaps 2;

Qy 3 MKQEQYVNTYPTKDGKVRIGDITDLWAEVHEHYTYTYGELKFGAGKTIREGQGN-SPD 61
Db 4 MSRKQYADMGPTVGDAIRLADSELFTIEIKDYTYTYGDEVKFGGKVRIDGQGPLATS 63
Qy 62 ENTLDLVITNAMIIDVTGIYKADIGIKNGIKHIGKAGNKMDDGVSPHVMVGVTEALA 121
Db 64 DECVDLVITNAMIIDVTGIYKADIGIKNGIKHIGKAGNPLMDGVD--MVIGAAATEVIA 121
Qy 122 GEGMIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGGTGPVDGNTNATTITPGKWNLRHM 181
Db 122 AEGMIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGGTGPVDGNTNATTITPGKWNLRHM 181
Qy 182 LRAAEYSNMVFLGKGNSSKKQLVQVEAGAIGFKLHEDWGTTPSAIDHCLSVADRYD 241
Db 182 LRAAEYSNMVFLGKGNSSKKQLVQVEAGAIGFKLHEDWGTTPSAIDHCLSVADRYD 241
Qy 242 VQVCHTDTVNEAGYVDDTLNANWGRAIHAYHIEGAGGSHSPDVTMAGELNLPSSSTP 301
Db 242 VQVAIHDTTDLNAGCVEDTMAAIAGRTHMTYHTEGAGGSHAPDIIVKAGEHNILPASTNP 301
Qy 302 TIPTYTINTVAEHLDMMLTCHLDKRIREDLOFSQSRIKRGSIKRYISKYTINPALTHGVSEY 361
Db 302 TIPTYTINTVAEHLDMMLTCHLDKRIREDLOFSQSRIKRGSIKRYISKYTINPALTHGVSEY 361
Qy 362 SOAMGRAGEVIRPTWQTADKNKKEFKLPEDGKNDNFRIKRYISKYTINPALTHGVSEY 421
Db 362 SOAMGRAGEVIRPTWQTADKNKKEFKLPEDGKNDNFRIKRYISKYTINPALTHGVSEY 421
Qy 422 IGSVEEGKIADLVVWNPAPFGVKPILVKGGMVVFSEMGDSNASVPTPQPVYRPMFGHH 481
Db 422 VGSVEVGKVLADLVWNPAPFGVKPILVKGGMVVFSEMGDSNASVPTPQPVYRPMFAAK 481
Qy 482 GKAKEDTSTITFVSKVAYENGVEKGLERQVLVKNCRNITKKDFKFNDKTAKITVDPKT 541
Db 482 GKAKEDTSTITFVSKVAYENGVEKGLERQVLVKNCRNITKKDFKFNDKTAKITVDPKT 541
Qy 542 FEVFDGKLCIT 552
Db 542 YHVFVDGKEVT 553

RESULT 15

US-09-252-991A-26887
; Sequence 26887, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26887
; LENGTH: 566
; TYPE: PRT

